Technical Support Document
EPA's 2011 National-scale Air Toxics Assessment

2011 NATA TSD

December 2015

Office of Air Quality, Planning, and Standards
Research Triangle Park
North Carolina 27711

This page intentionally left blank.

This page intentionally left blank.

This page intentionally left blank.

Contents

List of Exhibits

Common Acronyms and Abbreviations

µg/m^3	microgram/cubic meter
AERMOD	atmospheric dispersion model developed by the American Meteorological Society and the U.S. Environmental Protection Agency's Regulatory Model Improvement Committee
ASPEN	Assessment System for Population Exposure Nationwide
ATSDR	Agency for Toxic Substances and Disease Registry
CHAD	Consolidated Human Activity Database
CAP	Criteria air pollutant
CMAQ	Community Multiscale Air Quality model
CONUS	Continental United States (modeling domain for CMAQ)
EC	exposure concentration
EPA	Environmental Protection Agency
EGU	electricity generating unit
HAP	hazardous air pollutant
HAPEM	Hazardous Air Pollutant Exposure Model
HEM	Human Exposure Model
HI	hazard index
HQ	hazard quotient
IRIS	Integrated Risk Information System
ISC	Industrial Source Complex
MOVES	Motor Vehicle Emissions Simulator
NATA	National-scale Air Toxics Assessment
NEI	National Emissions Inventory
NMIM	National Mobile Inventory Model
OAQPS	Office of Air Quality Planning and Standards
PAH	polycyclic aromatic hydrocarbon
PM	particulate matter
POM	polycyclic organic matter
RfC	reference concentration
RTR	Risk and Technology Review
SCC	Source Classification Code
S/L/T	State, local, or tribal agency
URE	unit risk estimate
WRF	Weather Research Forecasting model

1 BACKGROUND AND INTRODUCTION

1.1 The Purpose of this Document

This document describes the data and approaches used to conduct the U.S. Environmental Protection Agency's (EPA; referred to throughout this document as "we") **National-scale Air Toxics Assessment** (NATA), an ongoing comprehensive evaluation of **air toxics** in the United States. It presents the approaches EPA used to conduct NATA, including descriptions of how we

- compiled emissions data and prepared them for use as model inputs,
- estimated ambient concentrations of air toxics,
- estimated exposures to air toxics for populations,
- selected toxicity values,
- characterized human-health risks and hazards, and
- addressed variability and uncertainty.

This technical support document (TSD) satisfies basic documentation protocol expected of EPA products and provides a resource for the technically oriented user community by summarizing the data sources, methods, models, and assumptions used in the 2011 NATA. References to additional documents are included (Section 8) to facilitate access to more detailed technical information on the **emissions inventories**, **dispersion modeling**, **photochemical modeling**, **exposure modeling**, and **toxicity values**.

Appendices to this document include:

- Appendix A—a glossary of the key terms and their definitions;
- Appendix B—a list of air toxics included in NATA and a list of source classification codes (SCCs) for diesel particulate matter (diesel PM);
- Appendix C—a crosswalk of pollutant names across inventories, assessments, and regulations, with metal speciation factors;
- Appendix D—a crosswalk table for NEI sectors to the source groups and Human Exposure Model (HEM-3) run groups used to present the NATA results, and additional speciation information including for xylenes, mercury, and other metals;
- Appendix E—procedures for estimating NATA **background concentrations**;
- Appendix F—additional **model evaluation summaries**;
- Appendix G—a table of average ratios of exposure concentration to ambient concentrations applied in NATA;
- Appendix H—a table of **toxicity values** applied in NATA; and
- Appendix I—adjustments to the approach.

We also provide a "SupplementalData" folder with this document that contains the Microsoft® Access™ and Microsoft® Excel™ files referenced throughout this TSD.

This document does not provide quantitative results for NATA and thus presents no exposure or risk estimates. Results and other specific information for NATA, including for the 2011 NATA and previous assessments, are found on the NATA website.

1.2 What NATA Is

NATA is a screening tool intended to evaluate the human-health risks posed by air toxics across the United States. We developed this tool so that state, local, and tribal agencies could prioritize air toxics, emission sources, and locations of interest for further study.

NATA assembles information on air toxics, characterizes emissions, and prioritizes air toxics and locations that merit more refined analysis and investigation. This information is used to plan, and assist with the implementation of, national, regional, and local efforts to reduce toxic air pollution. Using general information about sources to develop estimates of risks, NATA provides screening-level estimates of the risk of cancer and other potentially serious health effects as a result of inhaling air toxics. The resulting risk estimates are purposefully more likely to be overestimates of health impacts than underestimates, and thus they are **health protective**.

NATA uses emissions data compiled for a single year as inputs for modeling ambient air concentrations and estimating health risks. Results include estimates of ambient concentrations and exposure concentrations (ECs) of air toxics and estimates of **cancer risks** and potential **noncancer health effects** associated with chronic **inhalation** exposure to air toxics. The estimates are generated within each state, at both county and **census-tract** levels.

NATA provides a **"snapshot"** of outdoor air quality and the risks to human health that might result if air toxic emission levels were to remain at the same levels as those estimated for the assessment year. The estimates reflect only risks associated with chronic (relatively long-term) exposures to the inhalation of air toxics at the population level. The assumptions and methods used to complete the national-scale assessments limit the types of questions that NATA can answer reliably. These limitations, described throughout later sections of this document and summarized in Section 7, must be considered when **interpreting the NATA results** or when using them to address questions posed outside of NATA.

NATA results are useful for prioritizing air toxics and emission sources, identifying locations of interest that require additional investigation, providing a starting point for local-scale assessments, focusing community efforts to reduce local emissions of air toxics, and informing the design of new monitoring programs or the re-design of existing ones. NATA results also can provide general answers to questions about emissions, ambient air concentrations, and exposures and risks across broad geographic areas (such as counties, states, the nation) at a moment in time.

NATA was designed to answer questions such as the following:

- Which air toxics pose the greatest potential risk of cancer or adverse noncancer effects across the entire United States?
- Which air toxics pose the greatest potential risk of cancer or adverse noncancer effects in specific areas of the United States?

- Which air toxics pose less, but still significant, potential risk of cancer or adverse noncancer effects across the entire United States?
- When risks from inhalation exposures to all outdoor air toxics are considered in combination, how many people could experience a lifetime cancer risk greater than levels of concern (e.g., 1-in-1 million)?
- When potential adverse noncancer effects from long-term exposures to all outdoor air toxics are considered in combination for a given target organ or system, how many people could experience exposures that exceed the reference levels intended to protect against those effects (i.e., a hazard quotient greater than 1)?

1.3 The History of NATA

As discussed on the NATA website, EPA's first national-scale air toxics study was the Cumulative Exposure Project (Caldwell et al. 1998), which was developed based on estimates of air toxics emissions present before the Clean Air Act (CAA) was amended in 1990. The Cumulative Exposure Project provided estimates of outdoor air toxics concentrations in each of the more than 60,000 continental U.S. census tracts.

For the first NATA, the Cumulative Exposure Project framework was enhanced to include estimates of population exposure and health risk. The first NATA used a more refined inventory of air toxics emissions developed for 1996, known at that time as the **National Toxics Inventory**. This assessment was submitted for a technical peer review in January 2001 to a panel of EPA's **Science Advisory Board** (EPA 2001b). The panel provided detailed comments later that year on the validity of the overall approach, the elements of the assessment (including the data, models, and methods used), and the manner in which these components were integrated into a national-scale assessment (EPA 2001a). EPA incorporated many of the Science Advisory Board's suggestions into the assessment and published the results of that assessment in 2002. Since then, four assessments have been completed, based on national emission inventories that are updated significantly on a tri-annual basis, representative of air toxic emissions in 1999, 2002, 2005, and 2011, respectively. In general, the scope of NATA has progressively expanded with subsequent versions, and some methods have been refined and improved. Exhibit 1 summarizes the **five NATAs** EPA has conducted to date.

Exhibit 1. Summary of the Five Completed NATAs

Inventory Year	Year Completed/ Published	Air Toxics Modeled ^{a,b}	Key Attributes
1996	2002	33—32 HAPs, focusing on those of concern in urban areas; plus diesel PM	• ASPEN used to model ambient concentrations • HAPEM4 used to model inhalation exposures
1999	2006	177—176 HAPs, including all those with chronic-health toxicity values at the time; plus diesel PM	• ASPEN used to model ambient concentrations • HAPEM5 used to model inhalation exposures • Doubled the number of emission sources covered compared to 1996 NATA
2002	2009	181—180 HAPs, including 4 with additional health information; plus diesel PM	• ASPEN and HEM (with ISC) used to model ambient concentrations • HAPEM5 used to model inhalation exposures

Inventory Year	Year Completed/ Published	Air Toxics Modeled ^{a,b}	Key Attributes
2005	2010	179—178 HAPs, for which emissions data and chronic-health toxicity values are available; plus diesel PM	• Emissions inventory updated to include recent information on industrial sources, residual-risk assessments, lead emissions from airports, and other sources • ASPEN and HEM-3 (with AERMOD, a more refined dispersion model) used to model ambient concentrations; HEM used for more source types than in 2002 • Exposure factors derived from 2002 NATA used to estimate inhalation exposures • CMAQ model (EPA 2015g) used to estimate **secondary formation** of acetaldehyde, acrolein, formaldehyde, and decay of 1,3-butadiene to acrolein
2011	2015	180—179 HAPs for which emissions data are; available plus diesel PM	• CMAQ and HEM-3 more fully integrated as a hybrid modeling system for about 40 HAPS and diesel PM to improve mass conservation. • HEM-3 with background for remaining HAPs (also for areas outside the U.S. continental CMAQ modeling domain) not covered by the hybrid approach • HAPEM7 to model inhalation exposures for a subset of air toxics and used to provide exposure factors for the remaining air toxics

[a] Note that "air toxics" and "HAPs" are sometimes used interchangeably. In this document, however, "air toxics" refers to HAPs plus diesel PM. HAPs are those air toxics which we are required to control under Section 112 of the 1990 CAA Amendments (EPA 2015n). Diesel PM is not a HAP but is carcinogenic to humans, although we have not yet developed a **unit risk estimate** for it. Given these concerns, the adverse noncancer effects of diesel PM are estimated in NATA (using an Integrated Risk Information System **reference concentration**) but its cancer risks are not estimated.

[b] The number of air toxics included in a NATA emission inventory can be slightly larger than the number of air toxics actually modeled. Some air toxics are not modeled because of uncertainty in the emissions numbers or in the ability to model air concentrations or health risk accurately. For example, asbestos is included in the inventory but not modeled and they are not included in the counts presented in this table.

Notes:
HAPs = hazardous air pollutants; diesel PM = diesel particulate matter; ASPEN = Assessment System for Population Exposure Nationwide; HAPEM4, HAPEM5, HAPEM7 = Hazardous Air Pollutant Exposure Model, version 4, 5, and 7; HEM = Human Exposure Model; CMAQ = Community Multiscale Air Quality model. ISC and AERMOD are Gaussian dispersion models.

1.4 How States and EPA Use NATA Results

NATA was designed as a screening assessment and functions as a tool to inform both national and more localized efforts to collect air toxics information, to characterize emissions, and to help prioritize air toxics and geographic areas of interest for more refined data collection and analyses.

Ultimately, NATA results are intended to focus resources on air toxics, locations, or populations that are associated with the greatest potential health risks. Thus, **the goal of NATA** is to identify those air toxics of greatest potential concern with regard to their contribution to population risk. The results are used to set priorities for the collection of additional air toxics information, including emissions and monitoring data. NATA was designed to help guide efforts to reduce toxic air pollution and to provide information that can be used to further the already significant emissions reductions achieved in the United States since 1990.

We use NATA to identify those air toxics and source sectors (e.g., stationary sources, mobile sources) having the highest exposures and health risks. The assessment results also help to identify geographic patterns and ranges of risks across the country. Specifically, **we use NATA results to**

- identify pollutants and industrial source categories of greatest concern,
- improve understanding of health risks posed by air toxics,
- help set priorities for the collection of additional information,
- set priorities for improving emission inventories,
- expand and prioritize EPA's network of air-toxics monitors,
- support communities in designing their own local assessments,
- enhance targeted risk-reduction activities, and
- provide a multiple-pollutant modeling framework linking air toxics to the Criteria Pollutant Program (EPA 2015d).

1.5 How NATA Results Should *Not* Be Used

As described in Section 1.2, NATA is a screening-level assessment that was designed to answer specific types of questions. The underlying assumptions of NATA and the methods limit the range of questions that can be answered reliably. NATA results should *not* be used independently to characterize or compare risk at local levels (e.g., between neighborhoods), nor should they be used to estimate exposure or health risks for individuals or groups within small geographic areas such as census blocks or to design control measures for specific emissions sources or pollutants.

NATA evaluations use emissions data for a single year as inputs to models that yield concentration and risk estimates. These estimates reflect chronic exposures. Given these characteristics, **NATA results should *not* be used for the following**:

- as a definitive means to pinpoint specific risk values within a census tract,
- to characterize or compare risks at local levels such as among neighborhoods,
- to characterize or compare risk among states,
- to examine trends from one NATA year to another,
- as the sole basis for developing risk reduction plans or regulations,
- as the sole basis for determining appropriate controls on specific sources or air toxics, or
- as the sole basis to quantify benefits of reduced air toxic emissions.

The limitations of the assessment methods prevent NATA from serving as a stand-alone tool. Furthermore, although results are reported at the census tract level, average risk estimates are far more uncertain at this level of spatial resolution than at the county or state level. For analysis of air toxics in smaller areas, such as census blocks or in a suspected "hotspot," other tools such as site-specific monitoring and local-scale assessments coupled with refined and localized data should be used.

These caveats are integral to the **proper interpretation of NATA results**. NATA results should be used to address *only* those questions for which the assessment methods are suited. Moreover, as noted above, NATA results from different assessment years generally should not be compared to each other. From one assessment to the next, EPA has improved its methodology and incorporated additional data that enhance the utility of the results, but compromise the ability to compare across periods. Specifically, each subsequent assessment has offered the following relative to the previous NATA:

- a better and more complete inventory of emission sources,
- an overall increase in the number of air toxics evaluated[1], and
- updated health data for use in risk characterization.

Successive improvements in methodology and improved data make comparing earlier assessments with later assessments inappropriate. Differences in emissions, ambient concentrations, or risks observed in the results of two assessments might be due either to improvement in the assessment methodology or to actual changes in emissions, populations, or other "real-life" characteristics.

NATA is not used solely as the source of information leading to regulations or guiding the enforcement of existing rules. Thus, even though some of the methods used to conduct NATA are similar to those used in air-related risk assessments conducted under the CAA mandate (such as residual risk assessments of HAP emissions from point sources, or assessments of exposures to criteria air pollutants (CAPs) for evaluations of National Ambient Air Quality Standards), NATA fundamentally differs from such assessments in that it is not a regulatory program.

1.6 The Risk Assessment Framework NATA Uses

The methods applied in conducting NATA are consistent with the general risk assessment framework used throughout EPA. This section provides background information on EPA's risk assessment framework and summarizes the NATA process. The analytical components of this process are then described in detail in subsequent sections.

EPA has published a series of guidelines (EPA 2015e) that establishes and explains the recommended methods for assessing human-health risks from environmental pollution. Included in this series are recommendations for carcinogen risk assessment, exposure assessment, chemical mixtures risk assessment, and other major EPA-wide risk assessment guidelines. In addition, EPA developed the three-volume Air Toxics Risk Assessment (ATRA) Reference Library (EPA 2004a,b; EPA 2006a) as a reference for those conducting air toxics risk assessments. This library provides information on the fundamental principles of risk-based assessment for air toxics, how to apply those principles in various settings, and strategies for reducing risk at the local level. EPA's guidelines and methods are consistent with the National Research Council's recommendations on conducting risk assessments (NRC 1983, 1994).

As described in more detail in these guidelines and documents, EPA's risk assessment process has three phases (Exhibit 2), the second of which has two parts.

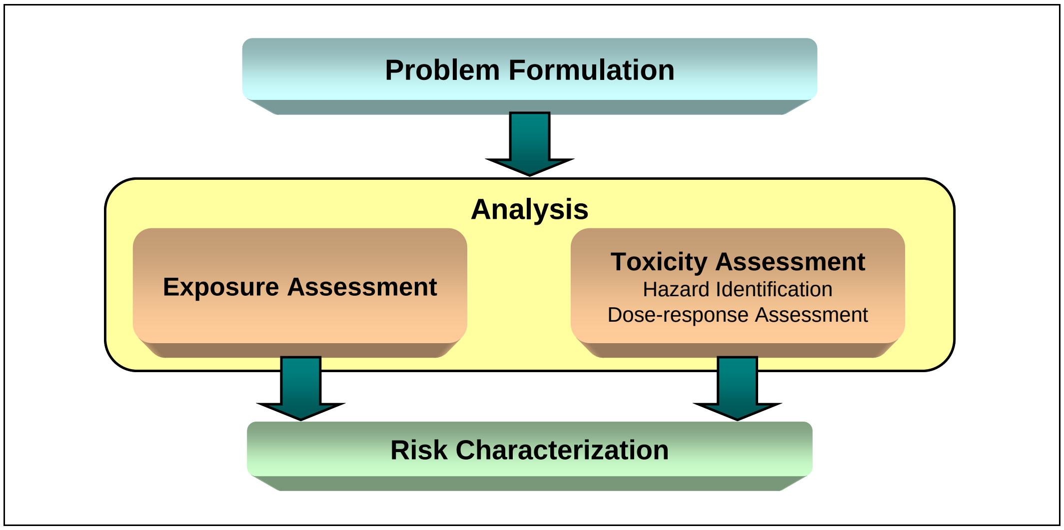
- The first phase (problem formulation) comprises the initial planning and scoping activities and definition of the problem, which results in the development of a conceptual model.

---

[1] Since the 1999 NATA, the number of pollutants has depended largely on the emissions inventory.

- The second phase (analysis) includes two components:
  - Exposure assessment; and
  - Toxicity assessment.
- The third phase is risk characterization, a synthesis of the outputs of the exposure and toxicity assessments to characterize health risks for the scenario described in the initial phase.

Exhibit 2. The General Air Toxics Risk Assessment Process



Source: Adapted from EPA (2004a)

An air toxics risk assessment starts with problem formulation. This initial step begins with the systematic planning and scoping that should be conducted before any analyses are begun to ensure that the objectives of the assessment are met, resources are used efficiently, and the overall effort is successful. One important product of the problem formulation for a risk assessment of air toxics is a conceptual model that describes how releases of air toxics might pose risks to people. The conceptual model serves as a guide or "road map" to the assessment. It defines the physical boundaries, potential sources and emitted air toxics, potentially exposed populations, chemical fate and transport processes, expected routes of exposure, and potential health effects.

This document is concerned primarily with describing the analysis phase of the general air toxics risk assessment process (and specifically with describing the analyses conducted for NATA). The analysis phase is the stage at which the risk assessment processes are used to evaluate the problem at hand. The planning and scoping activities and problem formulation we conduct before carrying out the analyses, however, are critical in that they set the course for the assessment and inform EPA's decisions regarding specific methods, models, and data sources to use. The conceptual model developed for NATA—which is the product of the first phase—is described in the following section. An overview of the analytical steps then follows in Section 1.8. Detailed descriptions of each step are presented in the other sections of this document.

1.7 The Scope of NATA

The national-scale assessment described in this document is consistent with EPA's definition of a cumulative risk assessment, as stated in EPA's Framework for Cumulative Risk Assessment (EPA 2003, p. 6), as "an analysis, characterization, and possible quantification of the combined risks to health or the environment from multiple agents or stressors." The Framework emphasizes that a conceptual model is an important output of the problem formulation phase of a cumulative risk assessment. The conceptual model defines the actual or predicted relationships among exposed individuals, populations, or ecosystems and the chemicals or stressors to which they might be exposed. Specifically, the conceptual model lays out the sources, stressors, environmental media, routes of exposure, receptors, and endpoints (i.e., measures of effects) relevant to the problem or situation that is being evaluated. This model takes the form of a written description and a visual representation of the relationships among these components. The conceptual model sometimes can include components that are not addressed specifically or quantitatively by an assessment, but that are nevertheless important to consider.

Section 2.4 of the report for the 1996 NATA presented to EPA's Science Advisory Board for review (EPA 2001b) included a conceptual model. Some of the specifics included in that conceptual model have since evolved as sequential assessments have been completed (for example, the number of air toxics evaluated has increased substantially since the 1996 NATA). The fundamental components included in NATA and the relationships among them, however, have been generally consistent for all five NATAs completed to date. Moreover, the conceptual model described in this document is very similar to the one presented in the documentation for the 1996 NATA.

NATA is national in scope, covering the United States, Puerto Rico, and the U.S. Virgin Islands. It focuses on **long-term inhalation exposures to air toxics**. In general, NATA is intended to provide EPA with the best possible national-scale population-level estimates of exposure to and risks associated with air toxics, taking into account data availability, technical capabilities, and other potentially limiting factors. The conceptual model for the 2011 NATA is presented in Exhibit 3. Each component included in the model is described briefly in the sections that follow.

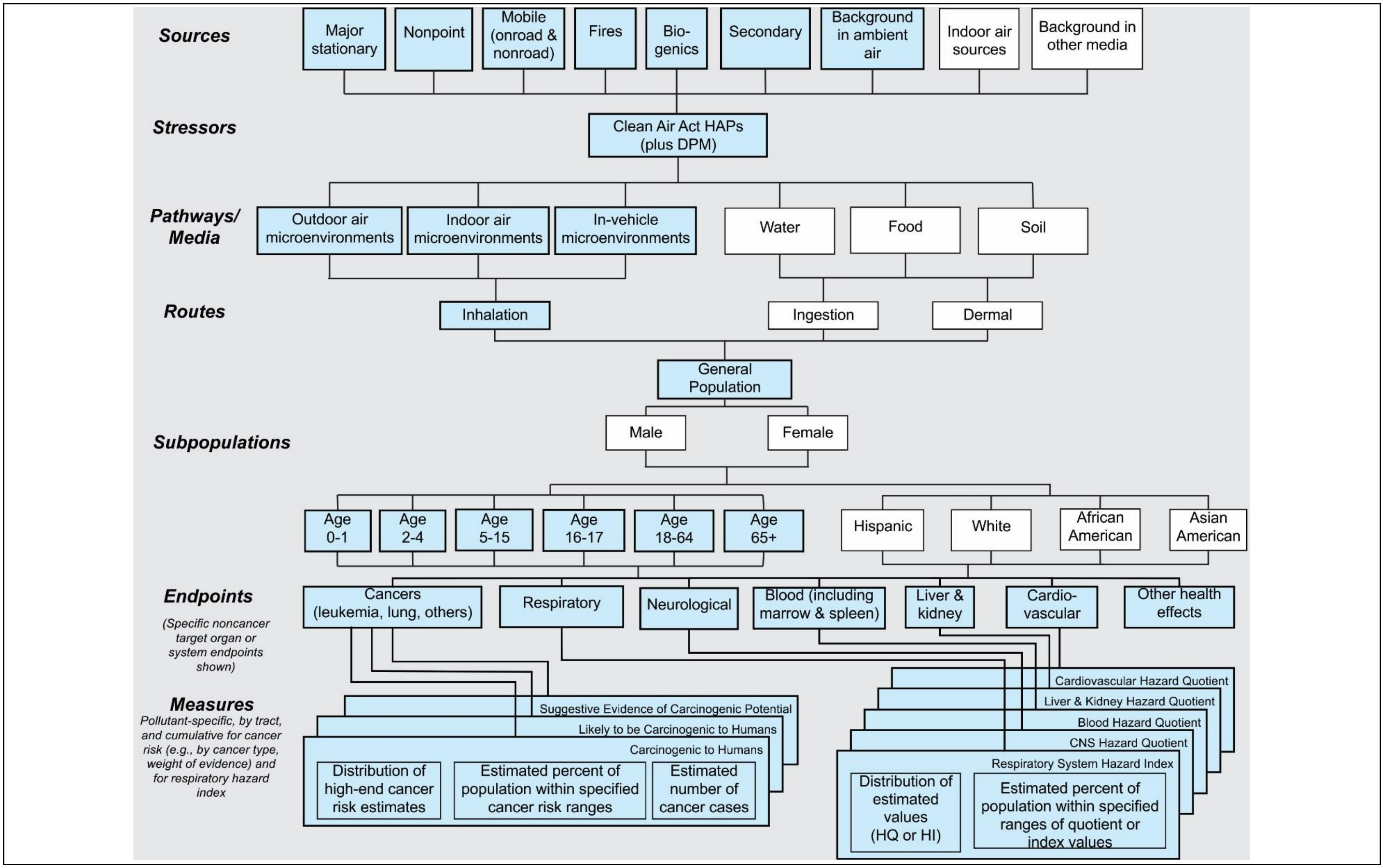
1.7.1 Sources of Air Toxic Emissions that NATA Addresses

Sources of primary air toxic emissions included in NATA (i.e., the NATA categories) are **point**, **nonpoint**, **mobile onroad and nonroad**, **biogenics**, and **fires** in the continental United States, and all these except biogenics and fires in Alaska, Hawaii, Puerto Rico and the U.S. Virgin Islands. Examples of point sources are large waste incinerators and factories. Nonpoint sources include residential wood combustion (RWC), commercial cooking, and consumer and commercial solvents. Mobile sources include vehicles found on roads and highways, such as cars and trucks, and nonroad equipment such as lawn mowers and construction equipment. Nonroad sources also include marine vessels, trains, and aircraft. Background sources, also included in NATA, can include natural sources and anthropogenic air toxics emitted in prior years that persist in the environment, or air toxics emitted from distant sources, including (for those HAPs modeled in HEM-3 but not the Community Multiscale Air Quality [CMAQ]) air toxics transported farther than 50 kilometers. Certain HAPs (i.e., formaldehyde, acetaldehyde, and acrolein) are formed in the atmosphere through photochemical reactions, and these "secondary" contributions are included in NATA through the photochemical air quality modeling platform. For the 2011 NATA, results are presented by broad categories and the more detailed NATA source groups through source attribution included in the air quality characterization. Details on the emission sources are presented in Section 2; details on air quality modeling and characterization are presented in Section 3.

1.7.2 Stressors that NATA Evaluates

The stressors evaluated through NATA can include any of the 187 HAPs defined in the 1990 CAA (190 HAPs were included originally but 3 have since been removed from the list). The set of air toxics included in NATA is determined by the emission and toxicity data available at the time of the assessment. **Diesel PM**, an indicator of diesel exhaust, is included in the set of stressors for NATA.

Exhibit 3. Conceptual Model for NATA



Blue boxes indicate elements included in the 2011 NATA; clear boxes indicate elements that could be included in future assessments. In the "Sources" included here, "Major stationary" includes both major and area sources as defined for regulatory purposes in the CAA. "Nonpoint" refers to smaller (and sometimes less discrete) sources that are typically estimated on a top-down basis (e.g., by county). Additional explanation of source types included in NATA is presented in Section 2. DPM refers to diesel PM. PBTs refers to chemicals that are persistent, bioaccumulative, and toxic. HQ and HI refer to hazard quotient and hazard index, respectively.

The 2011 NATA assesses the pollutants shown in Exhibit B-1 of Appendix B. Exhibit B-2 lists the CAA pollutants that are not included in the 2011 NATA and the reason. A spreadsheet file with more detailed information on the NEI and NATA pollutants is provided in the SupplementalData folder accompanying this TSD.

This exposure and risk assessment does not include the classes of compounds known as **dioxins**, **asbestos**, and **radionuclides**. We did not evaluate exposure and risk related to dioxins and radionuclides in the 2011 NATA because we did not evaluate the completeness or accuracy of the State, Local, and Tribal (S/L/T) agency data for these groups. Also, the most significant exposure route for dioxin is ingestion, not inhalation, so dioxin's relative contribution to NATA's inhalation risk estimates likely would not be large. Although the 2011 NATA emissions inventory includes asbestos, it also was not modeled for NATA because, like radionuclides, their ambient concentrations and inhalation exposures used in risk assessments typically are not expressed using mass-based concentrations, given methods used to develop the toxicity values that match each material's specific toxicological characteristics. Health risks of radionuclides are estimated using specific activity (a measure of radioactivity, which occurs as energy is emitted in the form of radiation from unstable atoms), and air concentrations of asbestos often are measured in terms of numbers of fibers per unit volume. The NEI currently is not compatible with emissions reported in units other than mass, and therefore suitable emissions data have not been compiled for these substances on a national scale.

1.7.3 Exposure Pathways, Routes, and Time Frames for NATA

Exposure to air toxics from all sources is determined by multiple interactions among complex factors, including the locations and nature of the emissions, the emission-release conditions, local meteorology, locations of receptor populations, and the specific behaviors and physiology of individuals in those populations. The particular combination of air toxics that people inhale, and the chemical interactions among those air toxics, influence the risks associated with these exposures. This high level of complexity makes aggregating risk across both substances and sources useful for depicting the magnitude of risks associated with inhalation of air toxics.

The air quality modeling step of NATA includes evaluating the transport of emitted particles and gases through the air to **receptors within 50 kilometers of sources**. Transformation of substances in the atmosphere (also referred to as secondary formation) and losses of substances from the air by deposition are included in the modeling, where data are available and the modeling approach supports it. For air toxics with sufficient ambient-monitoring data, or with emissions data primarily due to point sources, background concentrations are estimated. Taking into account fate and transport of emissions and the presence of some background concentrations, NATA estimates outdoor ambient concentrations across the nation.

NATA focuses on exposures due to inhalation of ambient air. Human receptors are modeled to account for an individual's movement among **microenvironments** such as residences, offices, schools, exterior work sites, and automobiles, where concentration levels can be quite different from general outdoor concentrations. The exposure assessment estimates air concentrations for each substance within each modeled microenvironment. The exposure assessment also accounts for human activities that can affect the magnitude of exposure (e.g., exercising, sleeping). This component of NATA accounts for the difference between ambient outdoor concentrations and the ECs (i.e., long-term-average concentrations to which people are actually exposed after taking into account human activities).

To date, NATA has not estimated air toxic concentrations in water, soil, or food associated with deposition from air, or the bioaccumulation of air toxics in tissues. Similarly, NATA has not estimated human exposures to chemicals via ingestion or dermal contact. EPA considers these pathways important but refined tools and data required to model multipathway concentrations and human exposures on the national scale are not yet readily available for use for many air toxics.

NATA estimates average annual outdoor concentrations that are used to develop long-term inhalation exposures for each of the air toxics. For cancer and chronic (long-term) health effects, the exposure duration is assumed as lifetime (i.e., 70 years for the purposes of this analysis). Subchronic and acute (lasting less than 24 hours) exposures are not estimated in NATA because the emissions database contains only annual-total emissions. If the emission inventories are later expanded to cover short-term (e.g., hourly, daily) emission rates, we would consider incorporating shorter exposure times into NATA.

1.7.4 Receptors that NATA Characterizes

NATA characterizes average risks to people belonging to distinct human subpopulations. The population as a whole is divided into **cohorts** based on residential location, life stage (age), and daily-activity pattern. A cohort is generally defined as a group of people within a population who are assumed to have identical exposures during a specified exposure period. Residential locations are specified according to U.S. Census tracts, which are geographic subdivisions of counties that vary in size but typically contain about 4,000 residents each. Life stages are stratified into **six age groups**: 0–1, 2–4, 5–15, 16–17, 18–64, and 65 and older. Daily-activity patterns specify time spent in various microenvironments (e.g., indoors at home, in vehicles, outdoors) at various times of day. For each combination of residential census tract and age, 30 sets of age-appropriate daily activity patterns are selected to represent the range of exposure conditions for residents of the tract. A population-weighted typical exposure estimate is calculated for each cohort, and this value is used to estimate representative risks, as well as the range, for a **"typical" individual residing in that tract**. Risk results for individual cohorts are not included in the outputs of NATA.

To date, NATA evaluations have **not included non-human receptors** (e.g., wildlife and native plants). The complexity of the varied ecosystems across the vast geographic area that is the scope of NATA precludes considering potential adverse ecological impacts at this time. Local- and urban-scale assessments could be developed to include non-human receptors, contingent on the availability of necessary resources, data, and methodologies. We currently, however, have no plans to include non-human receptors in NATA.

1.7.5 Endpoints and Measures: Results of NATA

NATA reports estimated cancer risks and noncancer hazards attributed to modeled sources. Key measures of **cancer risk** developed for the 2011 NATA include:

- upper-bound estimated lifetime individual cancer risk, and
- estimated numbers of people within specified risk ranges (e.g., number of individuals with estimated long-term cancer risk of 1-in-1 million or greater or less than 10-in-1 million).

For **noncancer effects**, the key measures presented in the 2011 NATA are **hazard indices** summed across all air toxics modeled for the respiratory system. Individual pollutant **hazard quotients** are provided for other target organs and systems.

NATA characterizes cancer risk and potential noncancer effects based on estimates of inhalation ECs determined at the census-tract level. This approach is used only to determine **geographic patterns of risks within counties**, and *not* to pinpoint specific risk values for each census tract. We are reasonably confident that the patterns (i.e., relatively higher levels of risk within a county) represent actual differences in overall average population risks within the county. EPA is less confident that the assessment pinpoints the *exact locations* where higher risks exist, or that the assessment captures the highest risks in a county. EPA provides the risk information at the census-tract level rather than just the county level, however, because the county results are less informative (in that they show a single risk number to represent each county). Information on variability of risk within each county would be lost if tract-level estimates were not provided. This approach is consistent with the purpose of NATA, which is to provide a means to inform both national and more localized efforts to collect air toxics information and to characterize emissions (e.g., to help prioritize air toxics and geographic areas of interest for more refined data collection such as monitoring). Nevertheless, the assumptions made in allocating mobile- and nonpoint source emissions within counties can result in significant uncertainty in estimating risk levels, even though general spatial patterns are reasonably accurate.

1.8 Model Design

Consistent with the general approach for air toxics risk assessment illustrated in Exhibit 2, the analysis phase of NATA includes two main components: estimating exposure and estimating toxicity. The outputs of these analyses are used in the third phase, risk characterization, which produces health-risk estimates that can be used to inform research or risk management. These two phases (analysis and risk characterization) represent the "core" of EPA's assessment activities associated with NATA. This set of activities is referred to here as the "NATA risk assessment process."

The NATA process can be characterized by four sequential components:

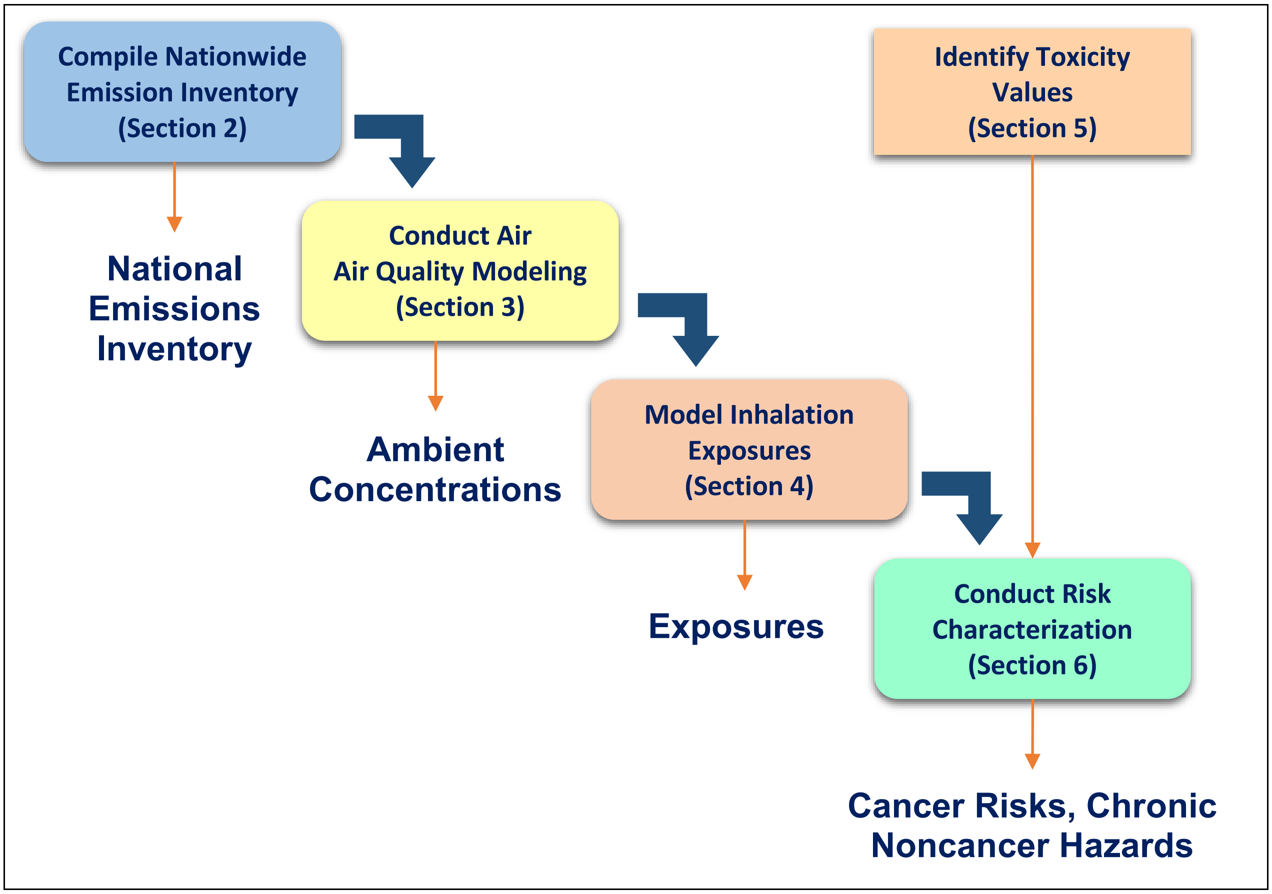
1. compiling the nationwide inventory of emissions from outdoor sources;
2. estimating ambient outdoor concentrations of the emitted air toxics across the nation;
3. estimating population exposures to these air toxics via inhalation; and
4. characterizing potential health risks associated with these inhalation exposures.

The fourth component (risk characterization) also requires that quantitative dose-response or other toxicity values be identified for each air toxic included in the assessment. These values are taken from those developed by other EPA and non-EPA programs. Although this step does not require a "new" quantitative dose-response assessment to be conducted as part of NATA, it does require that we make important scientific and policy decisions regarding the appropriate values to be used in NATA. Because these decisions are critical to the risk results, the identification of appropriate dose-response values is also described in this TSD as a fifth assessment component.

Collectively, these five components make up the NATA risk assessment process illustrated in Exhibit 4. The development of the emission inventory, air quality modeling, inhalation exposure modeling, and risk characterization must be conducted sequentially—the completion of each step requires outputs from the previous step, and toxicity values are required to carry out the risk-characterization calculations. Cancer risks and the potential for noncancer health effects are estimated using available information on health effects of air toxics, risk-assessment and risk-characterization guidelines, and estimated population exposures.

Each of these five components is described briefly here and explained in detail in the remainder of this document.

Exhibit 4. The NATA Risk Assessment Process and Corresponding Sections of this TSD



- Section 2 contains an explanation of the source types and air toxics included in the NATA emissions inventory. It also describes the processes we carried out to prepare the emissions for the air quality models.
- Section 3 contains a discussion of the models and procedures used to estimate ambient concentrations of air toxics, with links and references to technical manuals and other detailed documentation for the models used for NATA.
- Section 4 contains explanations of the processes used to estimate population-level exposure to outdoor ambient levels of air toxics, taking into account information on activities and other characteristics that can affect inhalation exposures.
- Section 5 contains descriptions of the dose-response values used for NATA, the sources from which these values are obtained, and assumptions made specific to NATA.
- Section 6 contains an overview of the calculations used to estimate cancer risk and potential noncancer hazard.
- Section 7 contains explanations of the uncertainties and limitations associated with the NATA process that must be considered when interpreting NATA results.

As noted at the beginning of this section, this document is intended to serve as a resource accompanying the most recent national-scale assessment—the 2011 NATA. Accordingly, although the following sections contain information on the NATA process that are generally applicable to all previous NATAs, references to specific technical processes and supporting details typically emphasize what was done for the 2011 NATA.

1.8.1 The Strengths and Weaknesses of the Model Design

EPA developed NATA to inform both national and more localized efforts to collect information and to characterize air toxics emissions (e.g., prioritize air toxics or geographic areas of interest for monitoring and community assessments). Because of this targeted objective, tools other than NATA might be more appropriate for assessing health risks outside the specific purpose of NATA (e.g., for evaluating risks from either a broader or more specific perspective). To further define and clarify what NATA should not be used for, this section contains descriptions of some of the important data and results that are *not* included in NATA.

- NATA does *not* include information that applies to specific locations. The assessment focuses on variations in air concentration, exposure, and risk among geographic areas such as census tracts, counties, and states. All questions asked, therefore, must focus on the variations among these geographic areas (census tracts, counties, etc.). Moreover, as previously mentioned, results are far more uncertain at the census-tract level than for larger geographic areas such as states or regions. (Section 7 contains discussions on the higher uncertainty at small geographic scales such as census tracts.) Additionally, NATA does not include data appropriate for addressing epidemiological questions such as the relationship between asthma or cancer risk and proximity of residences to point sources, roadways, and other sources of air toxics emissions.

- The results do *not* include impacts from sources in **Canada or Mexico** other than from limited pollutants and source groups. Thus, the results for states bordering these countries do not comprehensively reflect sources of transported emissions that could be significant.

- NATA does *not* include results for individuals. Within a census tract, all individuals are assigned the same ambient air concentration, chosen to represent a typical ambient air concentration. Similarly, the exposure assessment uses activity patterns that do not fully reflect the actual variations among individuals.

- The results do *not* include exposures and risk from all compounds. For example, of the 180 air toxics included in the 2011 NATA (some of which encompass multiple substances), only 138 air toxics have been assigned dose-response values. The remaining air toxics do not have adequate data in EPA's judgment to assess their impacts on health quantitatively, and, therefore, do not contribute to the aggregate cancer risk or target-organ-specific hazard indices. Of particular significance is that the assessment does not quantify cancer risk from diesel PM, although EPA has concluded that the general population is exposed to levels close to or overlapping with levels that have been linked to increased cancer risk in epidemiology studies. NATA, however, does model noncancer effects of diesel PM.

- Other than lead, which is both a CAP and a HAP, the results do *not* include the air pollutants, known as CAPs (particulate matter, ground-level ozone, carbon monoxide, sulfur oxides, nitrogen oxides), for which the CAA requires EPA to set National Ambient Air Quality Standards (other than CAP impacts on secondary formation of formaldehyde, acetaldehyde, and acrolein).

- The results do *not* reflect all pathways of potential exposure. The assessment includes risks only from direct inhalation of the emitted air toxics compounds. It does not consider air toxics compounds that might then deposit onto soil, water, and food and subsequently enter the body through ingestion or skin contact.

- The results do *not* include multipathway exposures because sufficiently refined tools and data required to model multipathway concentrations and human exposures for many air toxics on the national scale are not readily available for use.
- The assessment results reflect exposure at outdoor, indoor, and in-vehicle locations, but only to compounds released into the outdoor air, which could subsequently penetrate into buildings and vehicles. The assessment does *not* include exposure to air toxics emitted indoors, such as those from stoves, those that out-gas from building materials, or those from evaporative benzene emissions from cars in attached garages. The assessment also does not consider toxics released directly to water and soil.
- The assessment does *not* fully reflect variation in background ambient air concentrations. Background ambient air concentrations are average values over broad geographic regions.
- The assessment might *not* accurately capture sources that have episodic emissions (e.g., facilities with short-term deviations in emissions resulting from startups, shutdowns, malfunctions, and upsets). The models assume emission rates are uniform throughout the year.
- Short-term (acute) exposures and risks are *not* included in NATA.
- Atmospheric transformation and losses from the air by deposition are *not* accounted for in NATA air toxics that are not modeled in CMAQ.
- The evaluations to date have *not* assessed ecological effects, given the complexity of the varied ecosystems across the vast geographic area that NATA targets.

This page intentionally left blank.

2 EMISSIONS

The systematic compilation of a detailed, nationwide inventory of air toxics emissions is the first major step in the NATA risk assessment process. This section contains descriptions of the emissions used for the 2011 NATA. Section 2.1 contains summaries of the sources of emissions data included in NATA. Section 2.2 contains summaries of the processing of emissions for input into CMAQ (EPA 2015g), and Section 2.3 contains summaries of the processing for input into HEM-3 (see also the HEM-3 User's Guides, EPA 2014e).

For simplicity and consistency throughout this TSD, all aspects or details of the HEM-3 model are referred to overall as "HEM-3," although most often the AERMOD component of HEM-3 is pertinent to the discussion. EPA designed and maintains AERMOD separate and apart from HEM-3; HEM-3 merely incorporates AERMOD.

2.1 Sources of Emissions Data

NATA is intended to address outdoor emissions of all HAPs and diesel PM (together called "air toxics" in this document). To model air toxics, emissions of both air toxics and **CAPs** are used so that the chemical interactions that occur across all pollutants are addressed.

Sometimes **"air toxics" and "HAPs"** are used interchangeably. In this document, however, "air toxics" refers to the HAPs that EPA is required to control under Section 112 of the 1990 Clean Air Act (EPA 2015n) plus **diesel PM**. The 1990 Clean Air Act Amendments required EPA to control 190 HAPs (EPA 2008b) and provided for revisions to be made to that list. Currently, the list includes 187 HAPs. Diesel PM is not a HAP, and EPA currently does not have sufficient evidence to develop a unit risk estimate for it. Some evidence does indicate that localized high lifetime cancer risks are, however, associated with exposure to diesel PM. Given such concern, the potential adverse noncancer effects associated with diesel PM are estimated in NATA (using an Integrated Risk Information System reference concentration) but its cancer risks are not.

The 2011 NATA combines modeling from CMAQ and HEM-3 for the continental United States. CMAQ multipollutant modeling addresses all sources in the NEI for CAPs and about 40 HAPs. Emissions from outside the United States are represented by CMAQ boundary conditions for benzene, formaldehyde, and acetaldehyde. For the remaining "non-CMAQ" HAPs and non-CMAQ parts of the modeling domain (i.e., Alaska, Hawaii, Puerto Rico and the U.S. Virgin Islands), only HEM-3 is used. For these pollutants and geographic regions, spatially uniform background concentrations based on remote concentrations are added to the HEM-3-modeled data to represent influences from transport and emissions outside the modeling domain. Similar to previous NATAs, HEM-3 modeling addresses the 180 NATA HAPs and diesel PM included in NATA and all anthropogenic sources except prescribed and agricultural burning.

The main source of the emissions data for the CAPs and HAPs modeled for NATA is the **2011 NEI v2**. The NEI is a comprehensive and detailed estimate of air emissions of CAPs and HAPs from all air emissions sources in the United States, including the territories of Puerto Rico and the U.S. Virgin Islands, and offshore sources and commercial marine vessels (CMVs) in Federal Waters. The NEI is prepared every three years by the EPA based primarily upon emission estimates and emission model inputs provided by S/L/T air agencies for sources in their jurisdictions, and supplemented by data developed by the EPA. These data are submitted electronically to the Emissions Inventory System (EIS). CAPs must be submitted under the EPA's Air Emissions Reporting Requirements (AERR). HAPs are submitted voluntarily. Lead is both a HAP and a CAP, so it must be submitted under the AERR. For the 2011 NEI, facilities with potential to emit greater than 5 tons per year (TPY) were required to report lead.

To build as complete an NEI as possible, and to ensure use of up-to-date emission factors from test programs resulting from regulatory development, we gap-filled emissions using various sources of information including the Mercury and Air Toxics Rule (MATS) test data, the Toxics Release Inventory (TRI; EPA 2015j), and the application of HAP-to-CAP emission-factor ratios to CAP emissions reported by S/L/T.

Exhibit 5 contains a summary of the sources of emissions data in the NEI. More information on these data sources can be found in the 2011 NEI documentation.

Exhibit 5. NEI Data Sources for HAP Emissions

Stationary Point	Most stationary point source HAP data were submitted voluntarily by S/L/T. For electric generating units (EGUs), we estimated emissions using test-based or average emission factors from the 2010 test program conducted in support of MATS for metal HAPs, hydrochloric acid, and hydrogen fluoride (except where S/L/T data were from testing or the configuration of the units changed such that the MATS test results would no longer be applicable for 2011 emissions). For some point sources, EPA gap-filled HAPs. Sources of data included: rule-based emission factors (e.g., mercury for electric arc furnaces), TRI data for 2011, augmentation of HAPs using emission-factor ratios (of HAP to CAP) applied to S/L/T-reported CAP emissions, and 2008 emissions data for selected categories. The data sources for point-source data are provided in the NATA facility-level data emission summaries provided on the NATA website.
Point Airports	EPA estimates used the Federal Aviation Administration (FAA) Emission Dispersion Modeling System using landing and takeoff (LTO) information from FAA databases and updated with S/L/T inputs. For some airports (general aviation) without detailed aircraft-specific activity data, straight emission factors were used. Lead emissions were estimated based on per-LTO emissions factors, assumptions about lead content in the fuel, and lead-retention rates in the piston engines and oil. For some airports, estimates were provided by S/L/T. NEI has over 19,800 airports (including heliports and seaplanes).
Point Rail yards	EPA estimates were grown from the 2008 emissions that were developed by the Eastern Regional Technical Advisory Committee (ERTAC) to 2011 estimates. Locations were identified using a database from the Federal Railroad Administration. CAP emissions were estimated by applying emission factors to the total amount of distillate fuel used by locomotives. Each railroad company provided fleet mix information that allowed ERTAC to calculate railroad-specific emission factors. The company-specific, system wide fleet mix was used to calculate weighted average emissions factors for switchers operated by each Class I railroad. EPA emission factors were used for $PM_{2.5}$, SO_2, and NH_3. HAP emissions were estimated by applying toxic fractions to the VOC or PM estimates. For some rail yards, estimates were provided by S/L/T. The NEI has about 750 rail yards.
Stationary Nonpoint	Developed by EPA and/or submitted by S/L/T. Where S/L/T submitted CAPs but not HAPs, missing HAP emissions were augmented.
Biogenics	Based on Biogenic Emission Inventory System (BEIS3.60), using 2011 meteorology from the Weather Research Forecasting Model (WRF). Gridded emissions were used in NATA and summed to annual county-level estimates for the NEI. Includes VOC, NO_X, and 3 HAPs: formaldehyde, acetaldehyde, and methanol.
Locomotives	EPA estimates developed by applying growth factors to the 2008 NEI values based on railroad-freight-traffic data from the 2008 and 2011 R-1 reports submitted by all Class I rail lines to the Surface Transportation Board and employment statistics from the American Short Lines and Regional Railroad Association for Classes II and III. The emissions were allocated to line-haul-shape IDs and yard locations based on 2008 allocations. HAP emissions were estimated by applying toxic fractions to the VOC or PM estimates. For some areas, estimates were provided by S/L/T.

CMVs	Emissions from category 1 and category 2 (C1/C2) and category 3 (C3) marine vessels at ports or underway. C1/C2 includes fishing boats, ferries, tugboats, and vessels on the Great Lakes; C3 includes ocean going vessels and large ships. For C1/C2 marine diesel engines, the emission estimates were consistent with the 2011 Locomotive and Marine federal rule making. We derived HAP estimates by applying toxic fractions to VOC or PM estimates. These national emissions estimates were geographically allocated based on the available port and underway activity (e.g., Army Corps of Engineers Waterborne Commerce, National Marine Fisheries Service, etc.) C3 commercial marine inventories were developed for a base year of 2002 from gridded Emissions Control Area model data and then projected to 2011 by applying regional adjustment factors to account for growth. For some states, estimates developed by Lake Michigan Air Directors Consortium estimates replaced EPA's (these data include HAP emissions).
Onroad	We generated emissions using the October 2014 version of MOVES2014 (database version movesdb20141021; Motor Vehicle Emissions Simulator; EPA 2015k), then using SMOKE-MOVES to generate data for a detailed set of new source classification codes (SCCs), designed around emissions modes (SMOKE = Sparse Matrix Operator Kernel Emissions modeling system; Houyoux et al. 2000). These models use state- or EPA-provided input details, specific to each county. California's emissions were developed via their EMFAC onroad model, but VOC HAPs were speciated from California-reported VOC consistent with the MOVES2014 speciation, and SCCs were modified to match the rest of the country.
Nonroad	Except for California and Texas, data are from the National Mobile Inventory Model (NMIM; EPA 2015p), which uses the NONROAD model with 2011 state-submitted and/or default inputs. All metals are modeled in NMIM using emission factors, and all the other HAPs are estimated from VOC or PM by applying toxic fractions. California data are from the California Air Resources Board, which uses its own model, and Texas data are from the Texas Commission on Environmental Quality, which runs NONROAD (exception: Hg and arsenic are from NMIM).
Fires	We estimated agricultural burning (included in stationary nonpoint) using remote-sensing data, crop-usage maps, and emission factors. State data received from numerous states were used ahead of EPA estimates. Inventoried at the county level. Day- and location-specific prescribed burning and wildfires estimated via the SMARTFIRE2 system (which includes the BlueSky modeling framework) with inputs from State agencies where available. Georgia and North Carolina submitted estimates, Florida estimates were scaled to conform to Florida's summaries, and Delaware fires were modified (some fires zeroed out) per state comments.

2.1.1 Developing NATA Emissions from the 2011 NEI

Two modeling platforms were developed for the two air quality models run for NATA—CMAQ and HEM-3. The starting-point emission files for both were based primarily on "flat file" formats of the 2011 NEI produced by the EIS for the SMOKE modeling system. Onroad emissions were generated by SMOKE-MOVES (Motor Vehicle Emissions Simulator; EPA 2015k), and an FF10 summary of the data was developed for input into HEM-3.

The **grouping/speciation of NEI pollutants** for the purposes of NATA produced diesel PM from PM and produced various grouped pollutants from individual compounds reported in the NEI, such as grouping individual glycol ethers into the single NATA HAP "glycol ethers." Appendices B, C, D, and H, along with the spreadsheet file "NATA_Pollutants_AppendixB_AppendixC.xlsx" within the "SupplementalData" folder accompanying this TSD, show additional information on the NATA pollutants. All pollutants that were grouped for purposes of NATA are included in emission summaries as both the ungrouped NEI pollutant and the group sum. The group sum is called "Sum of" concatenated with the group name. All groups are listed below:

- Sum of Chromium VI (hexavalent) Compounds
- Sum of Cyanides
- Sum of Cresol/Cresylic Acid (Mixed Isomers)
- Sum of Glycol Ethers
- Sum of Nickel Compounds
- Sum of Polychlorinated Biphenyls (Aroclors)
- Sum of POMPAH
- Sum of Xylenes

The below subsections contain additional descriptions of these pollutants groupings.

2.1.1.1 Diesel PM

Diesel PM is neither a CAP nor HAP as defined by Section 112 of the CAA, and it is not a separate pollutant in the 2011 NEI. For NATA, we generated emissions of diesel PM using PM_{10} in the NEI from mobile-source engine-exhaust emissions for engines burning diesel or residual-oil fuels. Diesel PM emissions were set equal to PM_{10} emissions for these onroad and nonroad engines. Although stationary engines also can burn diesel fuel, only mobile-related diesel-engine SCCs were used. A list of the SCCs for which PM_{10} emissions were assigned to diesel PM and the corresponding NEI data category is provided in Exhibit B-3 of Appendix B.

2.1.1.2 PAH/POM

The PAH/POM results are presented as a group. The 2011 NEI v2, however, contains dozens of specific PAH/POM compounds with different unit risk estimates (UREs; i.e., the concentration that yields 1-in-1 million lifetime risk of cancer). For NATA, we grouped the individual PAH/POM compounds for modeling in CMAQ and HEM-3 based on the URE. These groups are shown in Exhibit 6 along with the specific NEI compounds (note: CAS = Chemical Abstracts Service). In NATA, the risk results are provided in the most aggregated form: PAHPOM. A spreadsheet version of this crosswalk is available in the SupplementalData folder ("NATA_Pollutants_AppendixB_AppendixC.xlsx").

Exhibit 6. 2011 NEI v2 PAHs Grouped for CMAQ and HEM-3 Modeling based on URE

CMAQ/ HEM-3 PAH Group	NEI Pollutant Code (CAS No.)	NEI Pollutant Description	NEI Category					URE 1/(μg/m³)
			Event	Nonpoint	Nonroad	Onroad	Point	
PAH_000E0	120127	Anthracene	✓	✓	✓	✓	✓	0
PAH_000E0	129000	Pyrene	✓	✓	✓	✓	✓	0
PAH_000E0	85018	Phenanthrene	✓	✓	✓	✓	✓	0
PAH_176E5	86748	Carbazole		✓			✓	1.76E-05
PAH_176E5	218019	Chrysene	✓	✓	✓	✓	✓	1.76E-05
PAH_880E5	130498292	PAH, total		✓	✓	✓	✓	0.000088

CMAQ/ HEM-3 PAH Group	NEI Pollutant Code (CAS No.)	NEI Pollutant Description	NEI Category					URE 1/(µg/m^3)
			Event	Nonpoint	Nonroad	Onroad	Point	
PAH_880E5	191242	Benzo[g,h,i,]Perylene	✓	✓	✓	✓	✓	0.000088
PAH_880E5	192972	Benzo[e]Pyrene	✓	✓			✓	0.000088
PAH_880E5	195197	Benzo(c)phenanthrene	✓					0.000088
PAH_880E5	198550	Perylene	✓	✓			✓	0.000088
PAH_880E5	206440	Fluoranthene	✓	✓	✓	✓	✓	0.000088
PAH_880E5	208968	Acenaphthylene		✓	✓	✓	✓	0.000088
PAH_880E5	2381217	1-Methylpyrene	✓					0.000088
PAH_880E5	2422799	12-Methylbenz(a)Anthracene		✓				0.000088
PAH_880E5	250	PAH/POM - Unspecified		✓			✓	0.000088
PAH_880E5	26914181	Methylanthracene	✓				✓	0.000088
PAH_880E5	65357699	Methylbenzopyrene	✓					0.000088
PAH_880E5	8007452	Coal Tar					✓	0.000088
PAH_880E5	832699	1-Methylphenanthrene		✓				0.000088
PAH_880E5	83329	Acenaphthene		✓	✓	✓	✓	0.000088
PAH_880E5	86737	Fluorene		✓	✓	✓	✓	0.000088
PAH_880E5	90120	1-Methylnaphthalene					✓	0.000088
PAH_880E5	91576	2-Methylnaphthalene		✓		✓	✓	0.000088
PAH_880E5	91587	2-Chloronaphthalene		✓			✓	0.000088
PAH_176E4	193395	Indeno[1,2,3-c,d]Pyrene	✓	✓	✓	✓	✓	0.000176
PAH_176E4	203123	Benzo(g,h,i)Fluoranthene		✓			✓	0.000176
PAH_176E4	203338	Benzo(a)Fluoranthene	✓	✓			✓	0.000176
PAH_880E5	284	Extractable Organic Matter (EOM)					✓	0.000088
PAH_176E4	205823	Benzo[j]fluoranthene					✓	0.000176
PAH_176E4	205992	Benzo[b]Fluoranthene		✓	✓	✓	✓	0.000176
PAH_176E4	207089	Benzo[k]Fluoranthene	✓	✓	✓	✓	✓	0.000176
PAH_176E4	224420	Dibenzo[a,j]Acridine					✓	0.000176
PAH_176E4	226368	Dibenz[a,h]Acridine					✓	0.000176
PAH_176E4	5522430	1-Nitropyrene					✓	0.000176
PAH_176E4	56553	Benz[a]Anthracene	✓	✓	✓	✓	✓	0.000176
PAH_176E4	56832736	Benzofluoranthenes	✓	✓			✓	0.000176
PAH_176E3	192654	Dibenzo[a,e]Pyrene					✓	0.00176

CMAQ/ HEM-3 PAH Group	NEI Pollutant Code (CAS No.)	NEI Pollutant Description	NEI Category					URE 1/(µg/m^3)
			Event	Nonpoint	Nonroad	Onroad	Point	
PAH_176E3	194592	7H-Dibenzo[c,g]carbazole					✓	0.00176
PAH_176E3	3697243	5-Methylchrysene		✓			✓	0.00176
PAH_176E3	41637905	Methylchrysene	✓					0.00176
PAH_176E3	50328	Benzo[a]Pyrene	✓	✓	✓	✓	✓	0.00176
PAH_192E3	53703	Dibenzo[a,h]Anthracene		✓	✓	✓	✓	0.00192
PAH_101E2	56495	3-Methylcholanthrene		✓			✓	0.01008
PAH_176E2	189559	Dibenzo[a,i]Pyrene					✓	0.0176
PAH_176E2	189640	Dibenzo[a,h]Pyrene					✓	0.0176
PAH_176E2	191300	Dibenzo[a,l]Pyrene					✓	0.0176
PAH_114E1	57976	7,12-Dimethylbenz[a]Anthracene		✓			✓	0.1136
PAH_880E5	779022	9-Methyl Anthracene		✓				0.000088

2.1.1.3 Metals

Example: Adjusting Emissions for Chromium VI Compounds

Chromic Acid (VI) (H_2CrO_4) has a molecular weight of about 118.01. Chromium, with an atomic mass of 52, is the toxic element of interest in this metal compound. Emissions reported in NEI are therefore multiplied by 0.4406 (i.e., 52 / 118.01), and the resulting emission rate is used in NATA modeling.

Metal emissions in the 2011 NEI represent only the mass of the metal or cyanide with a few exceptions for specific compounds of **hexavalent chromium** (chromium VI) and **nickel** of known composition (Exhibit 7). This is different from previous NATAs whereby the NEI included hundreds of specific metal compound (e.g., arsenic oxide). In such previous NATAs, after grouping or renaming the air toxics, emissions reported in NEI for each metal compound of known composition are adjusted so that the emissions rate used for NATA modeling corresponds to the mass of the elemental metal (or cyanide) only, and not the entire mass of the metal compound (see the adjacent text box for an example calculation). Thus, for the 2011 NEI these types of adjustments were made by S/L/T before submitting the emissions.

The three nickel compounds and three chromium VI compounds in the 2011 NEI are shown in the table below with the corresponding adjustment factors to compute the emissions that account for just the metal portion of the compound. Note that after applying the adjustments, the chromium VI compounds are grouped into chromium VI and the nickel compounds are grouped into nickel. Although the different nickel compounds have different UREs, they are combined in CMAQ and hence a single URE needs to be used. We chose the URE for nickel subsulfide. The vast majority of nickel in the NEI is “nickel” (pollutant code 7440020).

While hydrogen cyanide and cyanide may both be reported to the NEI, neither is adjusted in this way. **Calcium cyanamide** is adjusted to cyanide to allow the cyanide URE to be applied to only the cyanide portion of calcium cyanamide.

Exhibit 7. 2011 NEI Compounds or Compound Groups for which Emissions are Adjusted for CMAQ and HEM-3 Modeling

NEI Pollutant Code (CAS No.)	NEI Pollutant Description	NEI Category					NATA Website Pollutant Name	Metal/CN Speciation Factor
		Event	Nonpoint	Nonroad	Onroad	Point		
18540299	Chromium (VI)		✓	✓	✓	✓	CHROMIUM VI (HEXAVALENT)	1
7738945	Chromic Acid (VI)					✓	CHROMIUM VI (HEXAVALENT)	0.4406
1333820	Chromium Trioxide					✓	CHROMIUM VI (HEXAVALENT)	0.52
7440020	Nickel		✓	✓	✓	✓	NICKEL COMPOUNDS	1
1313991	Nickel Oxide					✓	NICKEL COMPOUNDS	0.7412
604	Nickel Refinery Dust					✓	NICKEL COMPOUNDS	1
156627	Calcium cyanamide						CALCIUM CYANAMIDE	0.4406

2.1.1.4 Other HAP Pollutant Groups: Xylenes, Cresols, and Glycol Ethers

Other HAPs may be reported to the NEI as compounds but are grouped together for purposes of NATA. These are glycol ethers, xylenes, cyanides, polychlorinated biphenyls, and cresols. The pre-grouped PAHs are further summed into the group PAHPOM. No multiplication is used to convert any individual compound in one of these groups; the mass is simply summed. Appendix C and a spreadsheet file ("NATA_Pollutants_AppendixB_AppendixC.xlsx") provided in the SupplementalData folder shows the specific NEI pollutants that constitute the groups.

2.1.2 Categorization of the NATA Emissions in the NATA Output Data

As explained on the NEI website, the 2011 NEI includes five data categories: point, nonpoint (formerly called "stationary area"), nonroad mobile, onroad mobile, and events consisting of wild and prescribed fires. NEI summaries are generally provided by sectors and tiers, which describe the type of emission source (e.g., industrial processes – oil and gas production). Some sectors and tiers cut across data categories since stationary sources are inventoried as both point and nonpoint. For example, the category "Fuel Comb - Comm/Institutional – Oil" results from large institutions inventoried as point sources (e.g., large universities with onsite steam plants) as well as commercial/institutional entities that are small and ubiquitous in nature, so their emissions are inventoried as county sums.

NATA summaries are provided by NATA broad summary categories and by more detailed source groups. The broad **NATA summary categories** are **point**, **nonpoint**, **onroad**, **nonroad**, **fires**, **biogenics**, and **secondary**. Some of these categories are named the same as the NEI data categories but they are not identical. For example, the NATA nonpoint category is not the same as the NEI nonpoint category because the NEI nonpoint category includes CMVs and locomotives, while the NATA nonpoint category does not. As another example, the NEI nonroad category is not the same as the NATA nonroad category because the NATA nonroad category includes airports, CMVs, and locomotives, while the NEI Nonroad category does not. Exhibit 8 contains comparisons between the NEI data categories and the NATA categories. Secondary is not included in Exhibit 8 since it not a primary emissions category covered in the NEI, but rather a result of atmospheric chemistry from the modeled emissions of CAPS and HAPs.

Exhibit 8. Map of NEI Data Categories to NATA Categories

NEI Data Category	NATA Category (Reflecting NATA Summary Results)
Point	**Point** [a]
Emissions estimates for sources that are individually inventoried and usually located at a fixed, stationary location, although portable sources such as some asphalt- or rock-crushing operations are also included. Point sources include large industrial facilities and electric power plants, but also increasingly include many smaller industrial and commercial facilities, such as dry cleaners and gas stations, which had traditionally been included as nonpoint sources. The choice of whether these smaller sources are estimated individually and included as point sources or inventoried as a nonpoint source aggregated to county or tribal areas is determined by the separate S/L/T air agency. The 2011 NEI v2 contains over 96,000 facilities (excluding about 1500 portable facilities that are not used in modeling due to a lack of geographic coordinates or county location); About 19,800 of these are airports.	Same as NEI point except: 1. Excludes portable sources, which are not modeled in either CMAQ or HEM-3 because no geographic information other than the state code is included. 2. Excludes airports and railyards, which are nonroad mobile.
Nonpoint	**Nonpoint**
Sources which individually are too small in magnitude or too numerous to inventory as individual point sources, and which can often be estimated more accurately as a single aggregate source for a county or tribal area. Examples are residential heating and consumer-solvent use. Agricultural, CMVs, and locomotive emissions are included. Biogenic emissions, which come from vegetation are also included but are estimated solely through EPA models.	Same as NEI nonpoint except excludes locomotives, CMVs, biogenic emissions, and agricultural fires
Onroad	**Onroad**
Estimates for mobile sources such as cars, trucks, and buses which are estimated via EPA models (other than in California which uses different models)—currently MOVES2014.	Same as NEI onroad.
Nonroad	**Nonroad**
Estimates for nonroad equipment such as lawn and garden equipment, and construction and recreational equipment which are typically estimated via EPA models (other than California which uses different models)—currently the NONROAD model which is run through NMIM. Note that emissions data for aircraft, locomotives, and CMVs are not included in NEI nonroad starting with the 2008 NEI.	Same as NEI nonroad but also including CMVs, locomotives and aircraft engine emissions occurring during LTOs, and the ground support equipment and auxiliary power units associated with the aircraft

NEI Data Category	NATA Category (Reflecting NATA Summary Results)
Event	**Fires**
Prescribed and wildfire emissions computed as day- and location-specific events	Wildfires, prescribed burning, and agricultural burning. These are modeled in CMAQ but not HEM-3. Wildfires and prescribed burning are generated via the SMARTFIRE2 model at specific geographic coordinates for each day, and are assigned to 12-km grid cells for input into CMAQ. Agricultural burning is inventoried at the county level and allocated to 12-km grid cells for input into CMAQ.
	Biogenic Emissions
	Emissions of formaldehyde, acetaldehyde, and methanol from vegetation (trees, plants, and soils) computed from the Biogenic Emission Inventory System within CMAQ. They are gridded to 12-km cells for CMAQ and are not modeled in HEM-3.

[a] In results presented online for assessments for the 2002 and earlier NATA inventories, point sources were divided into major sources and area sources and were sometimes referred to as stationary sources. Major sources are defined in the CAA as stationary sources that have the potential to emit either at least 10 TPY of a HAP or at least 25 TPY of any combination of HAPs. Area sources are stationary sources for which the locations are known but that emit at levels below the major source emissions thresholds. This terminology is not used in the 2011 NATA, and stationary-source emissions are referred to only as point-source or nonpoint-source emissions. Point sources in the NATA results refer to those sources, including smaller sources, for which a specific location for their emissions is identified by latitude and longitude descriptions, and nonpoint sources are those stationary sources that are not point sources.

2.1.3 Modifications to NEI Emissions Data

Although the 2011 NEI v2 is the main basis of the emissions fed into the air quality models, there were differences between the 2011 NEI v2 and emissions used for the NATA modeling. There were also differences between the emissions used for CMAQ and the emissions used for HEM-3 (see Section 2.1.4). Because CMAQ could not be rerun before the final NATA, all **emissions changes resulting from the 2011 NEI v2 review** were corrected via adjustments to HEM-3 and their impact on risk or adjustments to modeled concentrations and/or risks based on scaling. The resultant NATA emissions summarized with the results or provided with the maps reflects the adjustments made to HEM-3 modeling (except biogenics and fire emissions in the Continental United States, which were only run in CMAQ[2]).

The review of the 2011 NEI v2 data for NATA resulted in changes to emissions values and, for point sources, changes to geographic coordinates and release-point parameters. Nonpoint and onroad changes resulting from comments occurring after CMAQ modeling are provided in Exhibit 9, which also documents changes between the HEM-3 and CMAQ emissions. Other differences that result from differences in **emissions processing** (which reflect the specific role and function of the resulting inventory within the context of the NATA risk assessment process) are more accurately described as post-processing procedures rather than substantive changes. Examples of these types of changes include:

- For onroad emissions (CMAQ and HEM-3 used the same emissions although temporal allocation differed somewhat), changes were due to post-processing. Differences are provided in the spreadsheet file "2011eg NATA onroad versus NEIv2.xlsx" in the SupplementalData folder. The most significant is that manganese was higher (23.4 TPY in NATA versus 5 TPY in the NEI)

---

[2] Biogenic and fire emission summaries on the NATA website reflect only the emissions input into CMAQ.

because NATA includes manganese from brake and tire wear whereas the NEI does not. Manganese from brake and tire wear was computed from speciation of $PM_{2.5}$. There were also small differences in the HAP VOCs and PAHs in California due to changes in the post-processing approach to adjust California-submitted pollutants consistent with the MOVES2014 speciation. While both the NEI and NATA applied these adjustments, NATA used updated temporal profiles for California (which are included in the approach), resulting in slightly different emissions. Finally, the NATA inventory includes a more refined set of SCCs that includes road type to support spatial allocation of county-level emissions to finer scales.

- Emissions from the 2011 NEI for wildfires, prescribed burning, and agricultural-field burning are not used in NATA for Alaska, Hawaii and Puerto Rico (the U.S. Virgin Islands have none in the NEI) because these areas are not part of the CMAQ domain and these sources were only modeled in CMAQ.
- FIPS ending in 777 (inflight lead) were removed from the nonpoint inventory.
- Nonpoint tribal data (FIPS beginning in 88) were not used in the modeling because no surrogates are available and possible double counting would introduce uncertainty.
- For the residential wood-combustion sector, we removed emissions of 7,12-dibenzo(a,h)anthracene due to inconsistencies in estimates of this HAP between controlled and uncontrolled stoves (other than in the state of Minnesota in which 7,12- dibenzo(a,h)anthracene was retained in HEM-3 but not in CMAQ).
- For the residential wood-combustion sector, HAP-outlier issues were discovered with the data Clark County, Nevada submitted. The corrections made for the NATA emissions dropped some key pollutants. As a result, we replaced the NEI data with a complete set of CAPs and HAPs based on the EPA residential wood-combustion data set for Clark County.
- To minimize overstatement of fire emissions in single grid cells, we spatially allocated fires (from the event category, modeled only in CMAQ) that were larger than 20,000 acres on a single day. Those fires were projected as a circular area over the overlapping CMAQ grid cells. Each fire was then allocated to the grid cells it overlapped based on the area of overlap with that grid cell. In the emission inventories, these grid-cell "subfires" have names that end in _a, _b, etc. These are the identifiers of the 2011 fires that were split across grid cells:

State-County FIPS	Fire ID
56013	SF11C1791126
46065	SF11C1503125
48003	SF11C1718109
48243	SF11C1738273
48081	SF11C1742329
48415	SF11C1742358
48243	SF11C1747162
48125	SF11C1749358
48371	SF11C1750272
48353	SF11C1759082
35027	SF11C1760072
35027	SF11C1760460
32007	SF11C1774898
32013	SF11C1774993
32007	SF11C1775252

- Air toxic name conversions, placing individual air toxics into groups, and similar transcription and phraseology conversions (e.g., for the purpose of crosswalking the identity of an emitted air toxic to a substance with a quantitative dose-response value);
- Adjustments to emission rates of metal compounds based on the toxic-metal proportion of the compound's molecular weight (e.g., only the mass fraction of chromic acid that consists of chromium VI). In the emission summaries, pollutant groups entitled "Sum of….." includes the adjustments made, but emissions provided for the specific pollutants are presented pre-adjusted.

In addition, **background concentrations** for some air toxics are estimated for NATA based on monitoring and other data (no national inventory for emissions from background sources exists), and the **secondary formation** of a few air toxics is addressed in NATA but is not included in NEI.

2.1.4 2011 NATA Emissions: CMAQ versus HEM-3

Differences in the emissions inputs to CMAQ and HEM-3 were due to: (1) **design differences** in how the models were run and (2) differences due to timing of the model runs that did not allow for NATA review **comments from the S/L/T agency reviews** to be incorporated into the CMAQ modeling.

By **design**, there were differences in the sources of emissions used by the two models. The emissions input into HEM-3 excluded NATA categories more appropriately addressed by CMAQ: biogenics and three types of fires—wildfires, prescribed burning, and agricultural-field burning. Biogenic emissions were generated by running the Biogenic Emission Inventory System (BEIS) model in SMOKE using hourly meteorological inputs to generate hourly gridded (12 km x12 km) emissions of several photochemical-model species including three HAPs: formaldehyde, acetaldehyde, and methanol. This category is routinely part of CMAQ runs and is more appropriately modeled in CMAQ due to its broad spatial and refined temporal resolution and meteorological dependence. Wildfires and prescribed and agricultural burning were also run in CMAQ but not HEM-3 because CMAQ provides in-line plume rise to higher vertical layers based on the acres burned. Agricultural burning does not use these algorithms. We chose, however, to include agricultural burning with the other two types of fires as the spatial resolution for these estimates is more appropriate for CMAQ than the local-scale resolution provided by HEM-3. Also, grouping agricultural burning with the other fires enabled us to retain source attribution from the CMAQ zero-out runs (although not between the different fire types).

Also, CMAQ and HEM-3 treated emissions from outside the domain differently. CMAQ used boundary conditions for formaldehyde, benzene, and acetaldehyde, while other CMAQ pollutants used zero boundary conditions. For HEM-3, a spatially uniform remote concentration was added for non-CMAQ HAPs to all census-block centroids to account for transport and emissions coming from outside the domain. The secondary contribution (resulting from emissions from all categories that participate in photochemical reactions) was generated only in CMAQ, as HEM-3 does not account for secondary transformation.

In addition to differences in the sources of emissions used for in the two models, there were also differences in the way the emissions are processed. The processing of emissions for CMAQ and HEM-3 are described in Sections 2.2 and 2.3, respectively. Here we point out two differences in the resolutions of spatial and temporal allocation. For CMAQ, gridded (12 km x 12 km) and vertically allocated emissions were used for each modeled species for each hour of the year. Different temporal schemes were used for different categories to generate the hourly emissions, including continuous-emissions-monitoring data and hour- and pollutant-specific emissions for mobile sources, based on meteorologically dependent emission factors. For HEM-3, spatial and temporal approaches varied for different categories. Spatially, emissions were provided as point or HEM-3 "area" sources, where an area source could be (for example) a fugitive area within a facility, a runway at an airport, a port or underway water shape, or a census tract. This approach supports better resolution of air-concentration for sources with known locations (i.e., the sources in the NEI point category). Temporally, allocation schemes varied for different source categories. A summary of the spatial differences is provided in Exhibit 20 in Section 2.3.1.1.

For most categories, the resulting hourly emissions for HEM-3 and CMAQ used the same level of specificity, but some differences occurred in the nonpoint, onroad, and nonroad categories for which temporal profiles were aggregated from individual categories to a broader "HEM run group" (described in 2.3.2) and cannot vary by specific pollutant. Exhibit 21 in Section 2.3.1.2 contains descriptions of the temporal approach for each HEM run group.

In addition to design differences, there were also emission differences resulting from the **timing of the runs**. The emissions input to CMAQ did not include adjustments resulting from the review/comment periods. Below is a list of key differences in the emissions between CMAQ and HEM-3 due to comments during the NATA review/preview periods

Exhibit 9. Key Emission Differences between CMAQ and HEM-3 for 2011 NATA Modeling

Emissions category	CMAQ	HEM-3
Point	2011 NEI v2 with very small updates.	Post-NATA review and NATA-preview comments resulted in emission changes for about 200 facilities.
Diesel PM	Diesel PM was included in Federal Waters outside the U.S. boundaries (i.e., Gulf of Mexico) for C1/C2 marine but not for C3 vessels, Canada, or Mexico.	Diesel PM was included only in the United States. Two very small diesel PM sources in California were dropped because they had SCCs for railyards but not the facility type of the railyards (facility type was used for point-source attribution of airports and railyards).
Non-U.S. sources	Benzene, formaldehyde, acetaldehyde, and methanol are included from Canada, Mexico, and offshore sources.	Not included.
Onroad: Extended-idle emissions—California and Maricopa County	Used emissions consistent with 2011 NEI v2 / v6.2 platform.	For California—extended-idle emissions were changed to match the California Air Resources Board (CARB)-provided extended-idle emissions (RPH) by county/pollutant for CAPs, and split into extended-idle and APU emissions from v2. For HAPs, which CARB did not provide for RPH, we used CAP ratios to calculate HAP emissions. For example, if RPH VOC = 5% of total 220280 (heavy-duty diesel) VOC in a given county, then RPH benzene is set to 5% of total 220280 benzene in that county. All VOC HAPs (e.g., benzene, toluene) used VOC ratios. All PAHs (e.g., fluorene, benzo[a]pyrene) used $PM_{2.5}$ ratios. Pollutants with zero RPH emissions in SMOKE-MOVES, including all metals and dioxins/furans, were kept at zero. Maricopa extended-idle and APU emissions were cut in half consistent with the revised hoteling hours computed based on revised 2011 VMT by the *Maricopa Association of Governments*

Emissions category	CMAQ	HEM-3
Ethylene oxide (EO) from EO sterilizers	Used emissions consistent with 2011 NEI v2 (this pollutant is not in the v6.2 platform).	Removed EO sterilizer emissions submitted by the state of Maryland, which were based on 1999 methodology (prior to EO sterilization regulations). Note that this change was also made in Puerto Rico (we had carried forward older emissions in the 2011 NEI v2, and so we removed these for NATA).
Changes to trichloroethylene emissions in Clark County, Nevada	Used emissions consistent with 2011 NEI v2 (this pollutant is not in the v6.2 platform).	Replaced Clark County-submitted emissions with EPA estimates due to use of old EPA methodology.
Benzene from Utah counties—oil and gas	Used emissions consistent with 2011 NEI v2 / v6.2 platform.	Replaced benzene based on data provided by Utah (lower benzene to VOC).
Portable-fuel-container speciation impacting benzene and other HAPs in NY	Used emissions consistent with 2011 NEI v2 / v6.2 platform.	Replaced New York state emissions (all counties) with the EPA estimates that account for the benzene fuel limits from the MSAT rule.
7,12-dibenzo(a,h)anthracene emissions in MN from residential wood combustion	Removed from 2011 NEI v2 for purposes of v6.2 platform	Inadvertently did not remove this pollutant from the state of MN.

2.2 Emissions Preparation for CMAQ

EPA routinely prepares emissions for photochemical grid models through the development of an emissions modeling platform, and the **SMOKE modeling system** is used as the primary emissions modeling tool. An emissions modeling platform includes the emission inventories, the ancillary data files, and the approaches used to transform inventories for use in air quality modeling. Several platforms have been developed for 2011 NEI emissions. For NATA, EPA developed a multipollutant emissions modeling platform ("2011eg_nata_v6_11g 2011v2"), referred to here as the "**NATA CMAQ platform**," to generate the emission inputs for the version of CMAQ used for NATA (version MPv5.0.2). This version of CMAQ includes more air toxics than any other version—approximately 50 HAPs and diesel PM.

The NATA CMAQ platform is largely the same as was prepared for the 2011 NEI v2-based modeling platform (2011v6.2) with respect to the modeling domain (i.e., lower 48 states, parts of Canada and Mexico), grid resolution (12 km), and temporalization approaches. The platform differs in terms of speciation—it adds numerous air toxics model species required by CMAQ version MPv5.0.2.

2.2.1 Emission Inventories and Approaches: CMAQ

For the purposes of preparing the air quality model-ready emissions, the 2011NEIv2 was split into finer-grained sectors used for emissions modeling. The significance of an emissions modeling or "platform sector" is that the data are run through all of the SMOKE programs except the final merge (Mrggrid) independently from the other sectors. The sectors used for the NATA CMAQ platform were the same as used for the 2011 NEI v2 platform except that the point-source fires (prescribed and wild) were included as a single sector (ptfire). Exhibit 10 contains descriptions of the sectors used for the SMOKE/CMAQ modeling for NATA, and Exhibit 11 contains information on whether the sector contains HAP emissions.

Exhibit 10. Sectors Used in Emissions Modeling for the 2011 NATA CMAQ Platform

Platform Sector and Abbreviation	2011 NEI Data Category	Description and Resolution of the Data Input to SMOKE
EGU sector: *ptegu*	Point	2011 NEI v2 point-source EGUs. The 2011 NEI v2 emissions were replaced with hourly 2011 CEMS values for NO_X and SO_2, where the units were matched to the NEI. Other pollutants were scaled from 2011 NEI v2 using CEMS heat input. Emissions for all sources not matched to CEMS data came from 2011 NEI v2. Non-CEMS sources used daily resolution created using average fuel/region temporal files.
Point source oil and gas: *pt_oilgas*	Point	2011 NEI v2 point sources related to emissions from processes of oil and gas production based on the following North American Industry Classification System (NAICS) codes, though some individual facilities were moved to *ptnonipm* due to conflicting facility source types. **NAICS** **NAICS Description** 2111 Oil and Gas Extraction 2212 Natural Gas Distribution 4862 Pipeline Transportation of Natural Gas 21111 Oil and Gas Extraction 22121 Natural Gas Distribution 48611 Pipeline Transportation of Crude Oil 48621 Pipeline Transportation of Natural Gas 211111 Crude Petroleum and Natural Gas Extraction 211112 Natural Gas Liquid Extraction 213111 Drilling Oil and Gas Wells 213112 Support Activities for Oil and Gas Operations 221210 Natural Gas Distribution 486110 Pipeline Transportation of Crude Oil 486210 Pipeline Transportation of Natural Gas Annual resolution.
Remaining Non-EGU point: *ptnonipm*	Point	All 2011 NEI v2 point-source records not matched to the *ptegu* or *pt_oilgas* sectors, annual resolution. Includes all emissions from aircraft and ground-support equipment, which were inventoried at airports (point sources in the NEI), and some railyard emissions. Annual resolution.
Agricultural: *ag*	Nonpoint	NH_3 emissions from NEI nonpoint livestock and fertilizer application. County and annual resolution.
Area fugitive dust: *afdust*	Nonpoint	PM_{10} and $PM_{2.5}$ from fugitive-dust sources from the 2011 NEI v2 nonpoint inventory, including building construction, road construction, agricultural dust, and road dust. However, emissions from unpaved and paved road dust differ from the NEI in that the NEI data do not have a precipitation adjustment. Instead, the emissions modeling adjustment applies a transport fraction and a meteorology-based (precipitation and snow/ice cover) zero-out. County and annual resolution.
Nonpoint source oil and gas: *np_oilgas*	Nonpoint	2011 NEI v2 nonpoint sources from oil and gas-related processes. County and annual resolution
Residential wood combustion: *rwc*	Nonpoint	2011 NEI v2 NEI nonpoint sources with RWC processes. County and annual resolution.
Agricultural fires: *agfire*	Nonpoint	Agricultural burning from the 2011 NEI v2 nonpoint inventory. County and monthly resolution.

Platform Sector and Abbreviation	2011 NEI Data Category	Description and Resolution of the Data Input to SMOKE
Class 1 & 2 CMV and locomotives: *c1c2rail*	Nonpoint	Locomotives and primarily category 1 (C1) and category 2 (C2) sources of CMV emissions from the 2011 NEI v2 nonpoint inventory. County and annual resolution.
Commercial marine: *c3marine*	Nonpoint	Category 3 (C3) sources of CMV emissions from the 2011 NEI v2 nonpoint inventory. County and annual resolution.
Remaining nonpoint: *nonpt*	Nonpoint	2011 NEI v2 nonpoint sources not otherwise included in other platform sectors. County and annual resolution.
Nonroad: *nonroad*	Nonroad	2011 NEI v2 nonroad equipment emissions developed with NMIM using NONROAD2008 version NR08a. NMIM was used for all states except California and Texas, which submitted their own emissions to the 2011 NEI v2. County and monthly resolution.
Onroad: *onroad*	Onroad	2011 emissions of gasoline and diesel vehicles from parking lots and roadways. Includes the following modes: exhaust, extended idle, auxiliary-power units, evaporative, permeation, refueling, and brake and tire wear. For all states except California, based on monthly MOVES emissions tables from MOVES2014. California emissions are based on EMFAC. MOVES-based emissions computed for each hour and model grid cell using monthly and annual activity data (e.g., VMT, vehicle population). County and hourly resolution.
Onroad California: *onroad_ca_adj*	Onroad	2011 emissions of gasoline and diesel vehicles from parking lots and roadways, plus refueling, generated from SMOKE-MOVES, California only. Non-refueling California onroad emissions were adjusted so that they match emissions values provided by CARB, by county, by fuel type (non-diesel and diesel), and by the following vehicle types: motorcycles, passenger cars, and all other vehicles. County and annual resolution.
Point source fires: *ptfire*	Fires	Point-source day-specific wild and prescribed fires for 2011 computed using SMARTFIRE2 using State inputs where available, except for Georgia- and North Carolina-submitted emissions and corrections (scaling and zero-outs for certain fires) for Florida and Delaware. Consistent with 2011 NEI v2. Point and daily resolution.
Other point sources not from the NEI: *othpt*	N/A	Point sources from Canada's 2010 inventory and Mexico's 2008 INEM inventory. Also includes annual U.S. offshore oil 2011 NEI v2 NEI point-source emissions, and non-United States, non-Canada C3 CMV emissions. Annual resolution.
Other nonpoint and nonroad not from the NEI: *othar*	N/A	Annual year 2010 Canada (province resolution) and year 2008 Mexico INEM (municipio resolution) nonpoint and nonroad mobile inventories. Does not include Canadian *afdust* emissions. Annual resolution.
Other onroad sources not from the NEI: *othon*	N/A	Year 2010 Canada (province resolution) and year 2008 Mexico INEM (municipio resolution) onroad mobile inventories. Annual and province or municipio resolution.
Other area fugitive dust not from the NEI: *othafdust*	N/A	PM_{10} and $PM_{2.5}$ from fugitive dust sources from Canada's 2010 inventory. Emissions adjustments include a transport fraction and a meteorology-based (precipitation and snow/ice cover) zero-out. Annual and province or muncipio resolution.

Platform Sector and Abbreviation	2011 NEI Data Category	Description and Resolution of the Data Input to SMOKE
Biogenic: *beis*	N/A	Year 2011, hour-specific, grid cell-specific emissions generated from the BEIS3.6 model; includes emissions in Canada and Mexico. Hourly and grid-cell resolution.

Exhibit 11. Preparation of HAP Inventory for each Sector for the 2011 NATA CMAQ Platform

Platform Sector and Abbreviation	2011 NEI Data Category	Approach/Adjustments for HAPs
EGU sector: *ptegu*	Point	Used explicit HAPs from inventory. Did not speciate VOC to get primary HAP emissions.
Point source oil and gas: *pt_oilgas*	Point	Used explicit HAPs from inventory. Did not speciate VOC to get primary HAP emissions.
Remaining Non-EGU point: *ptnonipm*	Point	Used explicit HAPs from *ptegu* inventory. Did not speciate VOC to get primary HAP emissions. Generated diesel PM from PM_{10} emissions from diesel ground-support equipment and rail-yard emissions.
Agricultural: *ag*	Nonpoint	No HAPs in this sector.
Area fugitive dust: *afdust*	Nonpoint	No HAPs in this sector.
Nonpoint source oil and gas: *np_oilgas*	Nonpoint	Used explicit HAPs from inventory. Did not speciate VOC to get primary HAP emissions. Benzene overestimated in Utah in the 2011 NEI. Did not change CMAQ results but adjusted HEM-3 (and therefore hybrid).
Residential wood combustion: *rwc*	Nonpoint	Used explicit HAPs from inventory. Removed 7,12-dibenzo(a,h)anthracene due to inconsistency in different types of wood stoves.
Agricultural fire: *agfire*	Nonpoint	Used explicit HAPs from inventory.
Class 1 & 2 CMV and locomotives: *c1c2rail*	Nonpoint	Used explicit HAPs from inventory. Generated diesel PM from PM emissions of residual oil and diesel-fueled CMVs and locomotives.
Commercial marine: *c3marine*	Nonpoint	Used explicit HAPs from inventory. Generated diesel PM from PM emissions of residual oil and diesel-fueled CMVs.
Remaining nonpoint: *nonpt*	Nonpoint	Used explicit HAPs from inventory.
Nonroad: *nonroad*	Nonroad	Used explicit HAPs from inventory. Generated diesel PM from PM emissions of diesel-fueled engine-exhaust processes. Certain SCCs have no HAPs: national inventory does not have HAPs for liquefied petroleum gas and compressed natural gas (2267* and 2268*) and some records in the California inventory have VOCs but no HAPs. We did not add HAPs via speciation for situations in which there were no HAPs in the NEI.

<table>
<tr><th>Platform Sector and Abbreviation</th><th>2011 NEI Data Category</th><th>Approach/Adjustments for HAPs</th></tr>
<tr><td>Onroad: <i>onroad</i></td><td>Onroad</td><td>Used explicit HAPs from inventory.
MOVES integrates emissions of the following onroad-emitted species in development of the chemical-mechanism species such that the emissions of the explicit HAPs and chemical-mechanism species are consistent.
<table>
<tr><th>Pollutant ID</th><th>Pollutant Name</th><th>CMAQ HAP?</th></tr>
<tr><td>5</td><td>Methane (CH4)</td><td>n</td></tr>
<tr><td>20</td><td>Benzene</td><td>Y</td></tr>
<tr><td>21</td><td>Ethanol</td><td>n</td></tr>
<tr><td>22</td><td>MTBE</td><td>n</td></tr>
<tr><td>24</td><td>1,3-Butadiene</td><td>Y</td></tr>
<tr><td>25</td><td>Formaldehyde</td><td>Y</td></tr>
<tr><td>26</td><td>Acetaldehyde</td><td>Y</td></tr>
<tr><td>27</td><td>Acrolein</td><td>Y</td></tr>
<tr><td>40</td><td>2,2,4-Trimethylpentane</td><td>n</td></tr>
<tr><td>41</td><td>Ethyl Benzene</td><td>n</td></tr>
<tr><td>42</td><td>Hexane</td><td>n</td></tr>
<tr><td>43</td><td>Propionaldehyde</td><td>n</td></tr>
<tr><td>44</td><td>Styrene</td><td>n</td></tr>
<tr><td>45</td><td>Toluene</td><td>Y</td></tr>
<tr><td>46</td><td>Xylene</td><td>Y</td></tr>
<tr><td>185</td><td>Naphthalene gas</td><td>Y</td></tr>
</table>
Metal-HAP emissions were also included in this sector. Manganese, the only HAP included in $PM_{2.5}$ speciation, was estimated using a grams/mile emission factor for exhaust and from speciation of $PM_{2.5}$ for brake and tire wear.
Generated diesel PM from PM emissions of diesel-fueled engine-exhaust processes.</td></tr>
<tr><td>Onroad California: <i>onroad_ca</i></td><td>Onroad</td><td>HAPs were not exactly what California submitted to NEI. NEI used California's HAPs and total PAH. In the platform, we used California's VOCs as a starting point and speciated based on MOVES. We also speciated total PAH using specific-PAH-to-total-PAH ratios based on MOVES.</td></tr>
<tr><td>Point source fires: <i>ptfire</i></td><td>Fires</td><td>Used explicit HAPs from the inventory. Corrected CMAQ results for two PAH groups—PAH_176E3 was missing emissions (in some states) of methylchrysene due to use of incorrect pollutant code 248, and PAH_880E5 was missing emissions of methylbenzopyrene due us of incorrect pollutant code of 247.</td></tr>
<tr><td>Other point sources not from the NEI: <i>othpt</i></td><td>N/A</td><td>One of the Canada point inventories had pre-speciated VOCs, which included ALD2 (treated as acetaldehyde), FORM (formaldehyde), MEOH (methanol), but not BENZENE. This ALD2/FORM was mapped to ALD2_PRIMARY/FORM_PRIMARY in addition to itself, thus providing explicit HAP emissions for formaldehyde and acetaldehyde. In the other Canada point inventories and the Mexico point inventory, ALD2, FORM, ALD2_PRIMARY, FORM_PRIMARY, BENZENE, and MEOH were all created via speciation.</td></tr>
<tr><td>Other nonpoint and nonroad not from the NEI: <i>othar</i></td><td>N/A</td><td>No explicit HAPs. ALD2, FORM, ALD2_PRIMARY, FORM_PRIMARY, BENZENE, and MEOH were all created via VOC speciation.</td></tr>
<tr><td>Other onroad sources not from the NEI: <i>othon</i></td><td>N/A</td><td>No explicit HAPs. ALD2, FORM, ALD2_PRIMARY, FORM_PRIMARY, BENZENE, and MEOH were all created via VOC speciation.</td></tr>
</table>

Platform Sector and Abbreviation	2011 NEI Data Category	Approach/Adjustments for HAPs
Other area fugitive dust not from the NEI: *othafdust*		No HAPs in this sector.
Biogenic: *biog*	N/A	HAPs generated by BEIS included acetaldehyde (ALD2 and ALD2_PRIMARY), formaldehyde (FORM and FORM_PRIMARY), and methanol (MEOH).

2.2.1.1 Point Sources

The point-source components of the platform were derived from the SMOKE-formatted FF10 files exported from the emissions inventory system from September 2014. These data were supplemented with ethanol-plant emissions provided by EPA's Office of Transportation and Air Quality (OTAQ); these ethanol-plant emissions were included in 2011 NEI v2 but some had different names or EIS identifiers that were corrected in the HEM-3 data and NATA geoplatform. Point sources were separated into three sectors: ptegu, ptnonipm and pt_oilgas. The ptegu were separated due to the use of CEMs NO_X, SO_2, and heat input data for temporalization of NO_X, SO_2, and other pollutants to hourly. Other pollutants, including PM, VOCs, and HAPs, used hourly heat input. Airports were included in the ptnonipm sector.

2.2.1.2 Nonpoint Sources

The nonpoint-source components of the platform were derived from the SMOKE-formatted FF10 files exported from EIS. Biogenic emissions were removed since that sector is estimated from the BEIS model within SMOKE (see Section 2.2.1.6). The file was then split into rwc, c1c2rail, c3marine, np_oilgas, nonpt, afdust, ag, and agfire based on SCCs. Sectors afdust and ag had no HAPs.

The rwc sector included the SCCs shown in Exhibit 12; these are the same for the rwc modeling sector in the NATA CMAQ platform and the "RWC" HEM run group.

Exhibit 12. SCCs for RWC

SCC [a]	SCC Level 3	SCC Level 4
2104008100	Wood	Fireplace: general
2104008210	Wood	Woodstove: fireplace inserts; non-EPA certified
2104008220	Wood	Woodstove: fireplace inserts; EPA certified; non-catalytic
2104008230	Wood	Woodstove: fireplace inserts; EPA certified; catalytic
2104008310	Wood	Woodstove: freestanding, non-EPA certified
2104008320	Wood	Woodstove: freestanding, EPA certified, non-catalytic
2104008330	Wood	Woodstove: freestanding, EPA certified, catalytic
2104008400	Wood	Woodstove: pellet-fired, general (freestanding or FP insert)
2104008420	Wood	Woodstove: pellet-fired, EPA certified (freestanding or FP insert)
2104008510	Wood	Furnace: Indoor, cordwood-fired, non-EPA certified
2104008610	Wood	Hydronic heater: outdoor

SCC [a]	SCC Level 3	SCC Level 4
2104008700	Wood	Outdoor wood burning device, NEC (fire-pits, chimenas, etc.)
2104009000	Firelog	Total: All Combustor Types

[a] SCC levels 1 and 2 are Stationary Source Fuel Combustion and Residential.

The rwc temporalization was based on daily temperature (for day-specific emissions); different diurnal profiles were applied based on the SCC. More details are provided in the Technical Support document for the 2011 NEI v2-based Platform.

The c1c2rail and c3marine SCCs are shown in Exhibit 13, and the ag fire SCCs are shown in Exhibit 14.

Exhibit 13. SCCs for CMVs and Locomotive (c1c2rail and c3marine)

Sector	SCC [a]	SCC Level 2	SCC Level 3	SCC Level 4
c1c2rail	2280002100	Marine Vessels, Commercial	Diesel	Port emissions
c1c2rail	2280002200	Marine Vessels, Commercial	Diesel	Underway emissions
c1c2rail	2285002006	Railroad Equipment	Diesel	Line Haul Locomotives: Class I Operations
c1c2rail	2285002007	Railroad Equipment	Diesel	Line Haul Locomotives: Class II / III Operations
c1c2rail	2285002009	Railroad Equipment	Diesel	Line Haul Locomotives: Commuter Lines
c1c2rail	2285002008	Railroad Equipment	Diesel	Line Haul Locomotives: Passenger Trains (Amtrak)
c1c2rail	2285002010	Railroad Equipment	Diesel	Yard Locomotives
c3marine	2280003100	Marine Vessels, Commercial	Residual	Port emissions
c3marine	2280003200	Marine Vessels, Commercial	Residual	Underway emissions
c3marine	2280004000 [b]	Marine Vessels, Commercial	Gasoline	Total, All Vessel Types

[a] SCC level 1 for all of these is Mobile Sources.

[b] This SCC does not have any HAPs associated with it. It was reported by only Washington State (county level, not shape level) in the NEI.

Exhibit 14. SCCs for Agricultural-Field Burning (agfire)

SCC	Description
2801500000 [a]	* – whole field set on fire; Unspecified crop type and Burn Method
2801500600	* – whole field set on fire; Forest Residues Unspecified (see also 28–10–015–000)
2801500100	* – whole field set on fire; Field Crops Unspecified
2801500141	* – whole field set on fire; Field Crop is Bean (red): Headfire Burning
2801500170	* – whole field set on fire; Field Crop is Grasses: Burning Techniques Not Important
2801500181	* – whole field set on fire; Field Crop is Hay (wild): Headfire Burning

SCC	Description
2801500261	* – whole field set on fire; Field Crop is Wheat: Headfire Burning
2801500320	* – whole field set on fire; Orchard Crop is Apple
2801500330	* – whole field set on fire; Orchard Crop is Apricot
2801500350	* – whole field set on fire; Orchard Crop is Cherry
2801500390	* – whole field set on fire; Orchard Crop is Nectarine
2801500410	* – whole field set on fire; Orchard Crop is Peach
2801500420	* – whole field set on fire; Orchard Crop is Pear
2801500430	* – whole field set on fire; Orchard Crop is Prune
2801500500	* – whole field set on fire; Vine Crop Unspecified
2801500150	* – whole field set on fire; Field Crop is Corn: Burning Techniques Not Important
2801500220	* – whole field set on fire; Field Crop is Rice: Burning Techniques Not Significant
2801500250	* – whole field set on fire; Field Crop is Sugar Cane: Burning Techniques Not Significant
2801500262	* – whole field set on fire; Field Crop is Wheat: Backfire Burning
2801500300	* – whole field set on fire; Orchard Crop Unspecified
2801500440	* – whole field set on fire; Orchard Crop is Walnut
2801500450	* – whole field set on fire; Orchard Crop is Filbert (Hazelnut)

*=Miscellaneous Area Sources; Agriculture Production – Crops – as nonpoint; Agricultural Field Burning.

[a] This is the only SCC used for the EPA estimates; all other SCCs were used by state data submitters.

2.2.1.3 Onroad Sources

Highway-vehicle emissions data for NATA were largely based directly on the 2011 NEI and were generated using MOVES2014 (EPA 2015k) using Carbon Bond 05 (CB05) speciation.

SMOKE-MOVES uses the emission factors along with activity and meteorology data to produce hourly gridded emissions. This is explained in more detail in the Technical Support document for the 2011 NEI v2-based Platform.

For NATA, MOVES2014 was used everywhere except California. In California, we use the CARB-submitted data mapped to the MOVES2014 SCCs (new set of SCCs for onroad), except the VOC HAPs for which we used CARB VOCs speciated with MOVES2014-based profiles. These HAP VOCs are benzene, 1,3-butadiene, formaldehyde, acetaldehyde, naphthalene, acrolein, ethyl benzene, 2,2,4-trimethylpentane, hexane, propionaldehyde, styrene, toluene, xylene, and methyl-tert butyl ether. We chose this approach to achieve consistency in speciation and to use the current MOVES2014-based speciation data.

2.2.1.4 Nonroad Sources

The inventory estimates for nonroad engines were developed using the National Mobile Inventory Model (NMIM; EPA 2015p) for all sources except CMVs, locomotives, and aircraft, which are in different data categories of the NEI and are generated using different tools from NMIM. NMIM produces, in a consistent and automated way, county-level mobile-source emissions inventories nationwide for the NEI and for EPA rulemaking. NMIM is a consolidation of two EPA models: the Mobile Source Emission Factor model (MOBILE; EPA 2015o) and the NONROAD (EPA 2015q) model; but only the NONROAD portion was used for the 2011 NEI. NMIM estimates toxic emissions for nonroad sources using toxic-to-VOC ratios for gaseous air toxics, toxic-to-PM ratios for PAHs, and mass-per-mile emission factors for metals (EPA 2005b). MOVES2014a estimates toxic emissions from nonroad engines using updated data (EPA 2015b); however, this version of MOVES was not available at the time modeling for NATA 2011 was done.

EPA did not use NMIM for two states. California and Texas submitted emissions data that were used in place of EPA's NMIM run. Monthly emissions for these states were computed from the NMIM data. These states' data were further supplemented by EPA—for Texas, we added mercury and arsenic emissions from EPA's run of NMIM since they were missing from the Texas inventory; and, for California, missing VOC emissions from some SCCs were added as the sum of benzene, acetaldehyde, and formaldehyde.

The same data were used in 2011 NEI v2 as the 2011 NEI v1, other than for Delaware, California, and Texas. Delaware submitted updated inputs used for the 2011 NEI v2 and NATA.

2.2.1.5 Fires Sources: agfire and ptfire

Agricultural burning was pulled from the nonpoint FF10 file based on SCC codes. Prescribed fires and wildfires were already in their own sector. Corrections to ptfire emissions for two PAH groups were made after the CMAQ model was run, and so adjustments were applied to the CMAQ model concentrations.

The ptfire corrections were needed because we dropped pollutant codes 247 (methylbenzo(a)pyrene; assigned to PAH_880E5) and 248 (methylchrysene; assigned to PAH_176E3) from ptfire because these were incorrect codes in the ptfire modeling file for some states and we did not map that erroneous code to the PAH groups. The correct codes for these to HAPs are 65357699 (methylbenzo(a)pyrene) and 41637905 (methylchrysene). Delaware, Georgia, and North Carolina had correct codes and as did some fires in Nevada, New Mexico, South Dakota, Texas, and Wyoming. Correct codes were used in the EIS. Because other PAHs in ptfire are mapped to PAH_880E5 and PAH_176E3, we corrected concentrations for these PAH groups by multiplying the concentration by the ratio of the corrected PAH group's emissions to the erroneous PAH group's emissions. The ratios were computed at an annual state-specific resolution and applied to the grid cells based on an area-weighted average of state factors (for grid cells that were in more than one state).

A zero-out run of CMAQ provided source attribution from fires. In this zero-out run, emissions from both the agfires and ptfire sectors were set to zero.

2.2.1.6 Biogenics

Biogenic emissions were computed by running the BEIS3.60 model in SMOKE. BEIS creates gridded, hourly, model-species emissions from vegetation and soils. It estimates CO, speciated VOC, and NO emissions for the contiguous United States and for portions of Mexico and Canada. The speciated VOCs included isoprene, terpene, sesquiterpene, and three HAPs: formaldehyde, acetaldehyde, and methanol.

One of the updates we made was to be able to distinguish primary formaldehyde from secondary formaldehyde by assigning the formaldehyde to both FORM_PRIMARY and FORM species. This, along with a CMAQ zero-out run, provided a way to determine the source attribution of primary-emitted biogenic emissions. In this zero-out run, emissions from biogenic sources were set to zero.

In the United States, primary biogenic formaldehyde constitutes 71 percent of the total emissions of primary formaldehyde, primary biogenic acetaldehyde constitutes 86 percent of the total, and biogenic methanol constitutes 95 percent of the total.

2.2.2 Emissions Processing Steps and Ancillary Data

The processing steps and ancillary data associated with the spatial and temporal allocation were the same for the NATA platform as for the 2011 NEI v2 platform (used for other non-NATA modeling) and are documented at the 2011-based Modeling Platform at EPA's Clearinghouse for Inventories & Emissions Factors website. All sources were horizontally allocated to 12-km grid cells through either spatial surrogates or spatial coordinates. Sources were also vertically allocated to layer 1 or distributed vertically based on plume rise. A summary of these allocation methods is shown in Exhibit 15, and the Sections 2.2.2.1 and 2.2.2.2 expand further on this information.

Exhibit 15. Summary of Spatial and Temporal Allocation of Emissions for the 2011 NATA Platform

Platform Sector	Spatial-allocation Method	Speciation?	Temporal Resolution	Plume Rise
afdust	Surrogates	Yes	Annual	
ag	Surrogates	Yes	Annual	
agfire	Surrogates	Yes	monthly	
beis	Pre-gridded land use	in BEIS 3.60	computed hourly	
c1c2rail	Surrogates	Yes	annual	
c3marine	Surrogates	Yes	annual	
nonpt	Surrogates & area-to-point	Yes	annual	
nonroad	Surrogates & area-to-point	Yes	monthly	
np_oilgas	Surrogates	Yes	annual	
onroad	Surrogates	Yes	monthly activity, computed hourly	
onroad_ca_adj	Surrogates	Yes	monthly activity, computed hourly	
rwc	Surrogates	Yes	annual	
othafdust	Surrogates	Yes	annual	
othar	Surrogates	Yes	annual	
othon	Surrogates	Yes	annual	
othpt	Point	Yes	annual	in-line
pt_oilgas	Point	Yes	annual	in-line
ptegu	Point	Yes	daily & hourly	in-line
ptprescfire	Point	Yes	Daily	in-line

Platform Sector	Spatial-allocation Method	Speciation?	Temporal Resolution	Plume Rise
ptwildfire	Point	Yes	Daily	in-line
ptnonipm	Point	Yes	annual	in-line

2.2.2.1 Spatial Allocation

Sectors with county-level resolution were allocated to 12-km grid cells using spatial surrogates. Where data for a particular county were not available, a secondary or tertiary surrogate was used. Exhibit 16 contains the spatial surrogates available for the Continental United States (CONUS) domain used by CMAQ. These surrogates were assigned to sources based on SCC. Some county-specific SCC-to-surrogate assignments were made, as it was determined that at the census-tract level, the assignments of certain surrogates (particularly for RWC) were improperly concentrating emissions in some urban-county tracts. County-specific-surrogate assignments can be discerned from the first column of the cross-reference file noted below. The following files are available in the SupplementalData folder and provide more information on the surrogate assignments and underlying surrogate data:

- Spatial-surrogate cross-references by SCC ("SCC__spatialsurrogates_xref.xlsx")
- Spatial-surrogate definitions/sources of data ("US_SpatialSurrogate_Workbook_v072115.xlsx")
- "Technical Memorandum oil and gas surrogates and modeling data 2014-2-18.pdf" which describes the development of the oil and gas surrogates, surrogate codes 681-698

The 12-km spatial surrogate data are posted under the data for the 2011v6.2 Platform. Generally, the same spatial surrogate data (i.e., shapefiles of surrogate activity weights) were used for HEM-3 but they were computed separately from the underlying shapefiles or underlying raw data, as in the case of the updated oil and gas surrogates. As discussed in Section 2.3.1.1, the HEM-3 surrogates were allocated from counties to tracts instead of 12-km grid cells. Also, because the HEM-3 domain includes Alaska, Hawaii, Puerto Rico, and the U.S. Virgin Islands, some of the CONUS surrogates were not available and secondary surrogates had to be used. Furthermore, a few surrogate changes were made to HEM-3 after CMAQ was run—in particular, assignments from SCC to surrogate were made for specific counties, and three new county-specific surrogates were developed and used. These changes resulted from comments received during the NATA review, which revealed that, for some situations, that the surrogates were unreasonably concentrating emissions arbitrarily in one or two tracts in the county. Any county-specific surrogate can be identified using the first column of the surrogates-assignment file in the SupplementalData folder.

Exhibit 16. U.S. Surrogates Available for the 2011 Modeling Platform

Code	Surrogate Description
100	Population
110	Housing
120	Urban Population
130	Rural Population
137	Housing Change
140	Housing Change and Population
150	Residential Heating - Natural Gas
160	Residential Heating – Wood
165	0.5 Residential Heating - Wood plus 0.5 Low Intensity Residential
170	Residential Heating - Distillate Oil
180	Residential Heating – Coal
190	Residential Heating - LP Gas
200	Urban Primary Road Miles
210	Rural Primary Road Miles
220	Urban Secondary Road Miles
230	Rural Secondary Road Miles
240	Total Road Miles
250	Urban Primary plus Rural Primary
255	0.75 Total Roadway Miles plus 0.25 Population
260	Total Railroad Miles
270	Class 1 Railroad Miles
261	NTAD Total Railroad Density
271	NTAD Class 1, 2, 3 Railroad Density
280	Class 2 and 3 Railroad Miles
300	Low Intensity Residential
310	Total Agriculture
312	Orchards/Vineyards
320	Forest Land
330	Strip Mines/Quarries
340	Land

Code	Surrogate Description
520	Commercial plus Industrial plus Institutional
525	Golf Courses + Institutional +Industrial + Commercial
527	Single Family Residential
530	Residential - High Density
540	Retail Trade
545	Personal Repair
550	Retail Trade plus Personal Repair
555	Professional/Technical plus General Government
560	Hospital
565	Medical Office/Clinic
570	Heavy and High Tech Industrial
575	Light and High Tech Industrial
580	Food, Drug, Chemical Industrial
585	Metals and Minerals Industrial
590	Heavy Industrial
595	Light Industrial
596	Industrial plus Institutional plus Hospitals
600	Gas Stations
650	Refineries and Tank Farms
675	Refineries and Tank Farms and Gas Stations
680	Oil & Gas Wells, IHS Energy, Inc. and USGS
681-698	New set of Oil and Gas Surrogate Codes consistent with Oil and Gas Tool [a]
710	Airport Points
711	Airport Areas
720	Military Airports
800	Marine Ports
801	NEI Ports
802	NEI Shipping Lanes
807	Navigable Waterway Miles
808	Gulf Tug Zone Area

Code	Surrogate Description	Code	Surrogate Description
350	Water	810	Navigable Waterway Activity
400	Rural Land Area	812	Midwest Shipping Lanes
500	Commercial Land	850	Golf Courses
505	Industrial Land	860	Mines
510	Commercial plus Industrial	870	Wastewater Treatment Facilities
515	Commercial plus Institutional Land	880	Drycleaners
535	Residential + Commercial + Industrial + Institutional + Government	890	Commercial Timber

[a] See additional information in Exhibit 17 and "Technical Memorandum oil and gas surrogates and modeling data 2014-2-18.pdf" in the SupplementalData Folder.

Some of the surrogates listed in Exhibit 16 were not used in CMAQ and/or HEM-3. For example, the "Airport Areas" surrogate was used in HEM-3 but not CMAQ; most airport-related emissions were in the point inventory and did not need to be spatially allocated.[3] To look at the relative importance of the surrogates within the platform sectors, we computed the toxicity-weighted emissions for the CMAQ HAPs for the surrogates used in CMAQ (i.e., based on the SCC-to-surrogate assignments for CMAQ); these are shown in Exhibit 17.

---

[3] SCCs for airport aviation-gas distribution (2501080050, 2501080100, 25080201), which were in the nonpoint NEI category, uses code 711 ("airport areas") for HEM-3 but not CMAQ. For CMAQ, SMOKE used the area-to-point algorithm, which assigned these emissions to the geographic coordinates of airports using the ARTOPNT ancillary file, which included the geographic coordinates of major airports and allocation fractions to assign county-level emissions to multiple airports in the county.

Exhibit 17. Total and Toxicity-weighted Emissions of CMAQ HAPs Based on the CMAQ Surrogate Assignments

		Total CMAQ Emissions (HAP and Diesel PM): Fraction of Sector and Total									Cancer–weighted CMAQ Emissions: Fraction of Sector and Total									Respiratory–weighted CMAQ Emissions: Fraction of Sector and Total								
Surrogate Code	**Surrogate Description**	**agfire**	**c1c2rail**	**c3marine**	**nonpt**	**nonroad**	**np_oilgas**	**onroad**	**rwc**	**Total (TPY)**	**agfire**	**c1c2rail**	**c3marine**	**nonpt**	**nonroad**	**np_oilgas**	**onroad**	**rwc**	**Total (TPY)**	**agfire**	**c1c2rail**	**c3marine**	**nonpt**	**nonroad**	**np_oilgas**	**onroad**	**rwc**	**Total (TPY)**
100	Population				0.99	0.01				216,700				0.92	0.08				18,074				0.57	0.43				22,562
140	Housing Change and Population				0.11	0.89				60,448				0.07	0.93				16,629				0.00	1.00				170,238
150	Residential Heating – Natural Gas				1.00					591				1.00					2,960				1.00					1,543
165	0.5 Residential Heating – Wood plus 0.5 Low Intensity Residential								1.00	58,529								1.00	263,784								1.00	107,990
170	Residential Heating – Distillate Oil				1.00					99				1.00					5,490				1.00					315
180	Residential Heating – Coal				1.00					148				1.00					17				1.00					533
190	Residential Heating – LP Gas				1.00					49				1.00					248				1.00					173
200	Urban Primary Road Miles							1.00		59,697							1.00		63,561							1.00		123,133
205	Extended Idle Locations							1.00		21,888							1.00		59,215							1.00		70,100
210	Rural Primary Road Miles							1.00		37,604							1.00		34,492							1.00		93,523
221	Urban Unrestricted Roads							1.00		133,013							1.00		138,931							1.00		211,287
231	Rural Unrestricted Roads							1.00		81,752							1.00		83,453							1.00		153,654
240	Total Road Miles				1.00					2,557				1.00					261				1.00					133
250	Urban Primary plus Rural Primary				1.00					2,306				1.00					942				1.00					1
256	Off–Network Short–Haul Trucks							1.00		3,651							1.00		5,057							1.00		2,841
257	Off–Network Long–Haul Trucks							1.00		317							1.00		817							1.00		686
258	Intercity Bus Terminals							1.00		7							1.00		22							1.00		20
259	Transit Bus Terminals							1.00		24							1.00		70							1.00		63
260	Total Railroad Miles				1.00					58				1.00					24				0.00					0
261	NTAD Total Railroad Density		0.59			0.41				916		0.82			0.18				556		0.60			0.40				3,083

Surrogate Code	Surrogate Description	Total CMAQ Emissions (HAP and Diesel PM): Fraction of Sector and Total									Cancer–weighted CMAQ Emissions: Fraction of Sector and Total									Respiratory–weighted CMAQ Emissions: Fraction of Sector and Total								
		agfire	c1c2rail	c3marine	nonpt	nonroad	np_oilgas	onroad	rwc	Total (TPY)	agfire	c1c2rail	c3marine	nonpt	nonroad	np_oilgas	onroad	rwc	Total (TPY)	agfire	c1c2rail	c3marine	nonpt	nonroad	np_oilgas	onroad	rwc	Total (TPY)
271	NTAD Class 1 2 3 Railroad Density		1.00							28,137		1.00							15,144		1.00							98,881
280	Class 2 and 3 Railroad Miles		1.00							1,155		1.00							541		1.00							4,074
300	Low Intensity Residential				0.31	0.65			0.03	36,870				0.53	0.39			0.08	71,144				0.74	0.13			0.12	21,011
310	Total Agriculture	0.53			0.02	0.46				101,453	0.93			0.01	0.06				226,969	0.48			0.02	0.50				305,616
312	Orchards/Vineyards	0.97			0.03					539	0.61			0.39					3,502	1.00			0.00					1,424
320	Forest Land	0.99			0.01					97	1.00			0.00					391	0.99			0.01					264
330	Strip Mines/Quarries				1.00					9				1.00					302				1.00					6
340	Land					1.00				643					1.00				641					1.00				759
350	Water					1.00				87,373					1.00				53,031					1.00				8,205
400	Rural Land Area				0	1.00	0			119,240				0.00	1.00	0.00			43,368				0.00	1.00	0.00			7,442
500	Commercial Land				1.00					4,292				1.00					14,162				1.00					7,080
505	Industrial Land				0.74	0.26				21,496				0.84	0.16				15,594				0.35	0.65				25,782
506	Education							1.00		229							1.00		424							1.00		301
507	Heavy Light Construction Industrial Land							1.00		21							1.00		40							1.00		30
510	Commercial plus Industrial				0.11	0.31		0.59		82,646				0.02	0.30		0.67		86,727				0.17	0.49		0.34		61,363
515	Commercial plus Institutional Land				1.00					1,144				1.00					9,169				1.00					2,249
520	Commercial plus Industrial plus Institutional				0.04	0.96				44,805				0.00	1.00				42,735				0.00	1.00				13,253
526	Residential – Non-Institutional							1.00		473							1.00		352							1.00		128
527	Single Family Residential				1.00					9,681				1.00					2,103				1.00					1
535	Residential + Commercial + Industrial + Institutional + Government				0.13			0.87		345,828				0.00			1.00		319,891				0.05			0.95		112,380
540	Retail Trade (COM1)				1.00					48				1.00					15				1.00					29
545	Personal Repair (COM3)				1.00					11,879				1.00					1,019				1.00					780

Surrogate Code	Surrogate Description	Total CMAQ Emissions (HAP and Diesel PM): Fraction of Sector and Total									Cancer–weighted CMAQ Emissions: Fraction of Sector and Total									Respiratory–weighted CMAQ Emissions: Fraction of Sector and Total								
		agfire	c1c2rail	c3marine	nonpt	nonroad	np_oilgas	onroad	rwc	Total (TPY)	agfire	c1c2rail	c3marine	nonpt	nonroad	np_oilgas	onroad	rwc	Total (TPY)	agfire	c1c2rail	c3marine	nonpt	nonroad	np_oilgas	onroad	rwc	Total (TPY)
555	Professional/Technical (COM4) plus General Government (GOV1)				1.00					90				1.00					67				1.00					63
560	Hospital (COM6)				1.00					10				1.00					337				0.00					0
575	Light and High Tech Industrial (IND2 + IND5)				1.00					239				1.00					28				1.00					50
580	Food, Drug, Chemical Industrial (IND3)				1.00					554				1.00					1,777				1.00					490
585	Metals and Minerals Industrial (IND4)				1.00					10				1.00					4				1.00					1
590	Heavy Industrial (IND1)				1.00					25,345				1.00					22,024				1.00					291
595	Light Industrial (IND2)				1.00					21,569				1.00					885				1.00					987
596	Industrial plus Institutional plus Hospitals				1.00					0				0.00					0				0					0
600	Gas Stations				0.23			0.77		51,562				0.84			0.16		10,963				1.00			0.00		5
650	Refineries and Tank Farms				1.00					3,087				1.00					1,660				1.00					1
675	Refineries and Tank Farms and Gas Stations				1.00					16				1.00					9				0					0
680	Oil and Gas Wells						1.00			0						1.00			1						1.00			0
681	Spud count – Oil Wells						1.00			14						1.00			24						0			0
682	Spud count – Horizontally–drilled wells						1.00			70						1.00			202						1.00			201
683	Produced Water at all wells						1.00			107						1.00			202						0.00			0
684	Completions at Gas and CBM Wells						1.00			212						1.00			597						1.00			256
685	Completions at Oil Wells						1.00			309						1.00			835						1.00			328
686	Completions at all wells						1.00			1,053						1.00			2,434						1.00			1,321
687	Feet drilled at all wells						1.00			627						1.00			1,935						1.00			1,747
688	Spud count – Gas and CBM Wells						1.00			35						1.00			42						0.00			0

Surrogate Code	Surrogate Description	Total CMAQ Emissions (HAP and Diesel PM): Fraction of Sector and Total									Cancer–weighted CMAQ Emissions: Fraction of Sector and Total									Respiratory–weighted CMAQ Emissions: Fraction of Sector and Total								
		agfire	c1c2rail	c3marine	nonpt	nonroad	np_oilgas	onroad	rwc	Total (TPY)	agfire	c1c2rail	c3marine	nonpt	nonroad	np_oilgas	onroad	rwc	Total (TPY)	agfire	c1c2rail	c3marine	nonpt	nonroad	np_oilgas	onroad	rwc	Total (TPY)
689	Gas production at all wells						1.00			36,433						1.00			31,775						1.00			1,131
692	Spud count – all wells						1.00			254						1.00			709						1.00			707
693	Well count – all wells						1.00			835						1.00			1,895						1.00			663
694	Oil production at oil wells						1.00			8,720						1.00			15,982						1.00			1,648
695	Well count – oil wells						1.00			3,518						1.00			8,951						1.00			4,431
697	Oil production at gas and CBM wells						1.00			4,666						1.00			5,980						1.00			1,042
698	Well count – gas and CBM wells						1.00			13,023						1.00			40,438						1.00			28,683
700	Airport area				1.00					54				1.00					72				1.00					3
801	Port Areas				1.00					44				1.00					31				1.00					3
806	Offshore Shipping NEI2011 NOx		0.76	0.24						25,173		0.56	0.44						22,225		0.76	0.24						89,752
820	Ports NEI2011 NOx		0.34	0.66						6,620		0.08	0.92						17,240		0.34	0.66						23,643
850	Golf Courses					1.00				1,135					1.00				1,674					1.00				200
860	Mines					1.00				414					1.00				104					1.00				1,388
870	Wastewater Treatment Facilities				1.00					692				1.00					357				1.00					257
880	Drycleaners				1.00					8,300				1.00					830				1.00					0
890	Commercial Timber					1.00				2,263					1.00				1,313					1.00				3,233

Notes: Point sources including fires do not need to be spatially allocated using surrogates so are excluded from this table. Shading in the "Total" columns corresponds to the emission value, with warmer colors corresponding to higher emissions.

2.2.2.2 Speciation

Speciation converts the inventory species to the species needed by the model (model species). To generate the appropriate model species from the inventory species, inventory species need to be speciated and/or aggregated. These are done as a part of Smkinven and Spcmat SMOKE modules.

During the reading of the inventory, Smkinven uses the inventory table (ancillary file input to SMOKE) to subset the pollutants to only those that will be used in CMAQ. In addition, the inventory table provides the metal conversion factor for inventory compounds that contain additional non-metal components, to adjust the mass emissions to just the metal portion. Most metal-compound emissions in the NEI correspond to only the metal portion, so only a limited number of compounds require the metal conversion factor. Starting with the 2008 NEI and continuing for the 2011 NEI, most metal compounds (i.e., selenium oxide, lead nitrate) were retired (i.e., no longer allowed to be reported) and S/L/Ts were provided a set of factors to convert to just the metal portion. The few compounds reported for the 2011 NEI for which the SMOKE inventory table applies a metal conversion factor are nickel oxide (factor = 0.7412[4]), chromium trioxide (factor = 0.52), chromic acid (VI) (factor = 0.4406).

Also during the reading of the inventory, SMOKE computes NONHAPVOC by subtracting benzene, acetaldehyde, formaldehyde, and methanol (BAFM) from VOC for sources within sectors in which integration of CAP and HAP will occur (other than onroad since speciation is done within the MOVES2014 model). Integration allows the emissions of BAFM in the inventory to be used directly for speciation in place of the speciation fractions provided in the speciation profiles. This in turn provides consistency between the inventory and the CMAQ-model species for these explicit CMAQ HAPs. To implement HAP integration, the speciation profiles were revised to remove BAFM and were renormalized. The resultant non-BAFM profiles were renamed NONHAPTOG profiles. (TOG is total organic gases, and is VOC plus species of nonreactive organic gases such as methane.)

The remainder of the conversion of inventory species to model species occurs within Spcmat. Spcmat converts VOC and NONHAPVOC into model species using TOG and NONHAPTOG speciation profiles based on the inventory SCC. Spmat converts $PM_{2.5}$ into model species using PM speciation profiles also based on the SCC. Most inventory HAPs are mapped to themselves, except metals are speciated into coarse and fine particulates, and xylenes (mixed isomers) are speciated into m-, o- and p-xylenes. In addition, mercury is speciated into elemental, divalent, and particulate forms. We used unit-specific (i.e., specific to particular boiler unit at a power plant) profiles to speciate mercury from electricity generating units (EGUs) based on EGU speciation data. For new units since 2005, profiles were mapped to units based on the fuel, configuration, and technology. Profiles for other mercury-emitting categories are based on SCC, and were the same as those used in the 2005 platform (for mercury). Speciation profiles for mercury and other HAPs are in Appendix D.

Speciation with HAP integration is further described in detail in the 2011v6.2 platform documentation. Sources in nonpoint and nonroad data categories have partial integration. Sources need to have at least one BAFM compound to be integrated. For example, nonroad sources fueled with liquefied petroleum gas or compressed natural gas were not integrated because NMIM does not estimate any BAFM.

The one key difference between the 2011v6.2 platform and the NATA platform is the speciation of the non-integrated U.S. sources (i.e., point, fires, and any sources within nonpoint or nonroad that are not

---

[4] This factor should have been 0.7858 because the molecular weight for nickel is 58.6934 and for nickel oxide is 74.69. This occurred due possibly to an incorrect CAS for this compound.

integrated). For NATA, for non-integrated sources, we did not use explicit BAFM from the profiles, but rather we used these HAPs from the inventory. In this case, we created un-normalized NONHAPTOG profiles that remove BAFM, and we used inventory BAFM for the model species. In other platforms, the BAFM are dropped from non-integrated sources because these species are created from the inventory VOC.

The chemical mechanism used for the 2011 NATA platform was the CB05 mechanism (Yarwood 2005). The assignment of compounds to model species was based on a version 1 mapping of CB05 that assigned numerous compounds to "FORM" and "ALD2" (which were meant, in CB05, to be explicit formaldehyde and acetaldehyde). It was believed that these compounds reacted quickly in the atmosphere to form formaldehyde and acetaldehyde. However, subsequent to the NATA platform, these assignments were redone so that only formaldehyde and paraformaldehyde are assigned to FORM and only acetaldehyde is assigned to ALD2.

Exhibit 18 and Exhibit 19 contain the model species provided by SMOKE. The last column indicates whether the model species was added specifically for the NATA CMAQ Platform case. Note that some of the model species came from directly from the inventory and others were generated via speciation of inventory pollutants. As mentioned earlier, FORM, which is formaldehyde, was made up of inventory formaldehyde and some NONHAPVOC species (from speciation) that were mapped to FORM but were not explicit formaldehyde. Some inventory HAPs were mapped to multiple model species or were duplicated. For example, inventory formaldehyde was mapped to both FORM and FORM_PRIMARY, which allowed the secondary formaldehyde to be estimated (FORM – FORM_PRIMARY). Note FORM_PRIMARY was generated solely from inventory formaldehyde, and ALD2_PRIMARY was generated solely from inventory acetaldehyde. Inventory benzene was mapped to BENZENE, PAR, and UNR. Other than in Canada and Mexico, we did not create HAPs by speciating VOCs.

Exhibit 18. Gaseous Species Produced by SMOKE for the 2011 NATA Platform

Inventory Pollutant	Model Species	Description of Model Species	NATA Air Toxic?	Added for CMAQ NATA?
Cl_2	CL2	Atomic gas-phase chlorine	Y	N
HCl	HCL	Hydrogen Chloride (hydrochloric acid) gas	Y	N
CO	CO	Carbon monoxide	N	N
NO_X	NO	Nitrogen oxide	N	N
	NO2	Nitrogen dioxide	N	N
	HONO	Nitrous acid	N	N
SO_2	SO2	Sulfur dioxide	N	N
	SULF	Sulfuric acid vapor	N	N
NH_3	NH3	Ammonia	N	N
	NH3_FERT	Fertilizer ammonia	N	N
Benzene	BENZENE	Benzene	Y	N
Acetaldehyde	ALD2	Acetaldehyde	Y	N

Inventory Pollutant	Model Species	Description of Model Species	NATA Air Toxic?	Added for CMAQ NATA?
Acetaldehyde	ALD2_PRIMARY	Acetaldehyde-primary emissions	Y	Y
Acrolein	ACROLEIN	Acrolein	Y	N
Acrylonitrile	ACRYLONITRILE	Acrylonitrile	Y	Y
1,3-Butadiene	BUTADIENE13	1,3-Butadiene	Y	Y
Carbon tetrachloride[b]	CARBONTET[b]	Carbon tetrachloride[b]	Y	Y
Chloroform	CHCL3	Chloroform	Y	Y
1,4-Dichlorobenzene(p)	DICHLOROBENZENE	1,4-Dichlorobenzene(p)	Y	Y
1,3-Dichloropropene	DICHLORPROPENE	1,3-Dichloropropene	Y	Y
Ethylene dibromide (Dibromoethane)	BR2_C2_12	Ethylene dibromide	Y	Y
Ethylene dichloride (1,2-Dichloroethane)	CL2_C2_12	Ethylene dichloride	Y	Y
Ethylene oxide	ETOX	Ethylene oxide	Y	Y
Formaldehyde	FORM	Formaldehyde	Y	N
Formaldehyde	FORM_PRIMARY	Formaldehyde-primary emissions	Y	Y
Hexamethylene-1,6-diisocyanate	HEXAMETH_DIIS	Hexamethylene-1,6-diisocyanate	Y	Y
Hydrazine	HYDRAZINE	Hydrazine	Y	Y
Maleic Anyhydride	MAL_ANYHYDRIDE	Maleic Anyhydride	Y	Y
Methanol	MEOH	Methanol	Y	Y
Methylene chloride (Dichloromethane)	CL2_ME	Methylene chloride (Dichloromethane)	Y	Y
Naphthalene	NAPHTHALENE	Naphthalene	Y	Y
Specific PAHs assigned with URE = 0	PAH_000E0	Specific PAHs assigned with URE = 0	Y	Y
Specific PAHs assigned with URE = 1.76E-5	PAH_176E5	Specific PAHs assigned with URE = 1.76E-5	Y	Y
Specific PAHs assigned with URE = 8.80E-5	PAH_880E5	Specific PAHs assigned with URE = 8.80E-5	Y	Y
Specific PAHs assigned with URE = 1.76E-4	PAH_176E4	Specific PAHs assigned with URE = 1.76E-4	Y	Y
Specific PAHs assigned with URE = 1.76E-3	PAH_176E3	Specific PAHs assigned with URE = 1.76E-3	Y	Y
Specific PAHs assigned with URE = 1.76E-2	PAH_176E2	Specific PAHs assigned with URE = 1.76E-2	Y	Y

Inventory Pollutant	Model Species	Description of Model Species	NATA Air Toxic?	Added for CMAQ NATA?
Specific PAHs assigned with URE = 1.01E-2	PAH_101E2	Specific PAHs assigned with URE = 1.01E-2	Y	Y
Specific PAHs assigned with URE = 1.14E-1	PAH_114E1	Specific PAHs assigned with URE = 1.14E-1	Y	Y
Specific PAHs assigned with URE = 1.92E-3	PAH_192E3	Specific PAHs assigned with URE = 1.92E-3	Y	Y
Propylene dichloride (1,2-Dichloropropane)	PROPDICHLORIDE	Propylene dichloride (1,2-Dichloropropane)	Y	Y
Quinoline	QUINOLINE	Quinoline	Y	Y
1,1,2,2-Tetrachloroethane	CL4_ETHANE1122	1,1,2,2-Tetrachloroethane	Y	Y
Tetrachloroethylene (Perchloroethylene)	CL4_ETHE	Tetrachloroethylene (Perchloroethylene)	Y	Y
Toluene	TOLU	Toluene	Y	Y
2,4-Toluene diisocyanate	TOL_DIIS	2,4-Toluene diisocyanate	Y	Y
Trichloroethylene	CL3_ETHE	Trichloroethylene	Y	Y
Triethylamine	TRIETHYLAMINE	Triethylamine	Y	Y
m-xylene, xylene (mixed isomers)	MXYL	m-xylene, xylene (mixed isomers)	Y	Y
o-xylene, xylene (mixed isomers)	OXYL	o-xylene, xylene (mixed isomers)	Y	Y
p-xylene, xylene (mixed isomers)	PXYL	p-xylene, xylene (mixed isomers)	Y	Y
Vinyl chloride	CL_ETHE	Vinyl chloride	Y	Y
Mercury	HGNRVA	Elemental mercury	Y	Y
	HGIIGAS	Divalent gaseous mercury	Y	Y
VOC	ALD2	Acetaldehyde		N
	ALDX	Propionaldehyde and higher aldehydes		N
	CH4	Methane [a]		N
	ETH	Ethene		N
	ETHA	Ethane		N
	ETOH	Ethanol		N
	FORM	Formaldehyde		N
	IOLE	Internal olefin carbon bond (R-C=C-R)		N
	ISOP	Isoprene		N

Inventory Pollutant	Model Species	Description of Model Species	NATA Air Toxic?	Added for CMAQ NATA?
	MEOH	Methanol		N
	NVOL	Non-volatile		N
	OLE	Terminal olefin carbon bond (R-C=C)		N
	PAR	Paraffin carbon bond		N
	TERP	Terpenes		N
	TOL	Toluene and other monoalkyl aromatics		N
	UNK	Unknown		N
	UNR	Unreactive		N
	XYL	Xylene and other polyalkyl aromatics		N
VOC species from the biogenics model that do not map to model species above	SESQ	Sesquiterpenes		N
	NR	Non-reactive		N

[a] Technically, CH_4 is not a VOC but part of TOG.

[b] Carbon tetrachloride CMAQ concentrations were not used; it was dropped from the hybrid.

Exhibit 19. Particulate Species Produced by SMOKE for the 2011 NATA Platform

Inventory Pollutant	Model Species	Model species description	NATA Air Toxic?	Added Special for NATA Platform
PM_{10}	PMC	Coarse PM > 2.5 microns and ≤ 10 microns	N	N
	DIESEL_PMC	Coarse PM > 2.5 microns and ≤ 10 microns, diesel sources	Y	Y
$PM_{2.5}$	PEC	Particulate elemental carbon ≤ 2.5 microns	N	N
	PNO3	Particulate nitrate ≤ 2.5 microns	N	N
	POC	Particulate organic carbon (carbon only) ≤ 2 .5 microns	N	N
	PSO4	Particulate Sulfate ≤ 2.5 microns	N	N
	PAL	Particulate aluminum≤ 2.5 microns	N	N
	PCA	Particulate calcium ≤ 2.5 microns	N	N
	PCL	Particulate chloride ≤ 2.5 microns	N	N
	PFE	Particulate iron ≤ 2.5 microns	N	N
	PH2O	Particulate water ≤ 2.5 microns	N	N
	PK	Particulate potassium ≤ 2.5 microns	N	N

Inventory Pollutant	**Model Species**	**Model species description**	**NATA Air Toxic?**	**Added Special for NATA Platform**
	PMG	Particulate magnesium ≤ 2.5 microns	N	N
	PMN	Particulate manganese ≤ 2.5 microns	N	N
	PMOTHR	Other particulate matter ≤ 2.5 microns	N	N
	PNA	Particulate sodium ≤ 2.5 microns	N	N
	PNCOM	Particulate non-carbon organic matter ≤ 2.5 microns	N	N
	PNH4	Particulate ammonium ≤ 2.5 microns	N	N
	PSI	Particulate silica ≤ 2.5 microns	N	N
	PTI	Particulate titanium ≤ 2.5 microns	N	N
	DIESEL_PMEC	Particulate elemental carbon ≤ 2.5 microns, diesel sources	Y	Y
	DIESEL_PMFINE	Other particulate matter ≤ 2.5 microns, diesel sources	Y	Y
	DIESEL_PMNO3	Particulate nitrate ≤ 2.5 microns, diesel sources	Y	Y
	DIESEL_PMOC	Particulate organic carbon (carbon only) ≤ 2.5 microns, diesel sources	Y	Y
	DIESEL_PMSO4	Particulate Sulfate ≤ 2.5 microns, diesel sources	Y	Y
Nickel , nickel oxide, nickel refinery dust	NICKEL_C	Coarse particulate nickel	Y	Y
	NICKEL_F	Fine particulate nickel	Y	Y
Chromium VI, chromium (VI) acid, chromium trioxide	CHROMHEX_C	Coarse particulate chromium VI	Y	Y
	CHROMHEX_F	Fine particulate chromium VI	Y	Y
Chromium III	CHROMTRI_C	Coarse particulate chromium III	Y	Y
	CHROMTRI_F	Fine particulate chromium III	Y	Y
Arsenic	ARSENIC_C	Coarse particulate arsenic	Y	Y
	ARSENIC_F	Fine particulate arsenic	Y	Y
Beryllium	BERYLLIUM_C	Coarse particulate beryllium	Y	Y
	BERYLLIUM_F	Fine particulate beryllium	Y	Y
Cadmium	CADMIUM_C	Coarse particulate cadmium	Y	Y
	CADMIUM _F	Fine particulate cadmium	Y	Y
Manganese	MANGANESE_C	Coarse particulate manganese	Y	Y
	MANGANESE_F	Fine particulate manganese	Y	Y
Lead	LEAD_C	Coarse particulate lead	Y	Y
	LEAD_F	Fine particulate lead	Y	Y

Inventory Pollutant	Model Species	Model species description	NATA Air Toxic?	Added Special for NATA Platform
Mercury	PHGI	Particulate divalent mercury	Y	Y

2.3 Emissions Preparation for HEM-3

For the 2011 NATA, we used HEM-3 to run AERMOD. HEM-3 does not perform the emission-processing steps of temporal and spatial allocation, nor does it provide defaulted release characteristics, so there were a number of processing steps that needed to take place.

The key processing steps for preparing emissions for HEM-3 (i.e., for AERMOD) were source characterization temporal and spatial allocation. We developed "HEM run groups" to organize the modeling by these processing steps. Different HEM-3 runs were made for each grouping due to spatial, temporal, and release-parameter (height/building) differences.

It should be noted that the following emission sources were not modeled in HEM-3 (only modeled in CMAQ): agricultural burning, prescribed fires and wildfires (which were combined as "fires"), and biogenic emissions.

2.3.1 Overview of Differences in Emissions Processing Between CMAQ and HEM-3

Sections 2.3.1.1 and 2.3.1.2 below contain discussions on the spatial and temporal allocation used in the HEM-3 modeling and how they differ from those used in the CMAQ modeling.

2.3.1.1 Spatial Allocation

Exhibit 20 contains the different geometries used for inputting emissions into HEM-3 depending on the types of sources in the inventory. For CMAQ, all emissions for the 2011 NATA were gridded horizontally at 12-km resolution, regardless of the category. For HEM-3, different sources used different treatment. Point sources were modeled using the geographic coordinates provided in the inventory. For some facilities, multiple release points were given different coordinates to characterize the specific release points at a facility. The release-point-specific coordinates were used directly by HEM-3. For other facilities, there may be multiple release points provided, but only a single latitude and longitude. In this situation, all release points were modeled at that single location. Fugitive-release points at facilities were also associated with geographic coordinates that were used for modeling. Airports were similar; however, a separate dataset of runway information was assigned to NEI airports and used (details are provided in Section 2.3.4).

Exhibit 20. Approach for Spatial Allocation—HEM-3 versus CMAQ

Category	NEI Resolution	Spatial Approach for HEM-3	Spatial Approach for CMAQ
Point (excluding airports)	Point	Point—vertical stack and fugitive based on NEI information on emission-release point	12-km grid cells, vertical based on plume calculations
Airports	Point	Point—runways & 10-m-square areas consistent with NEI geographic coordinates	12-km grid cells

Category	NEI Resolution	Spatial Approach for HEM-3	Spatial Approach for CMAQ
Locomotives	Point (railyards) and County/Shape	Nonpoint—tracts Point—point fugitives	12-km grid cells
CMVs, ports, and underway	County/Shape	Shapes from the NEI; separate shapes used for CMV at ports versus underway	12-km grid cells
Onroad, nonroad equipment, and other nonpoint	County	Census tracts	12-km grid cells
Agricultural burning and biogenic emissions	County	Not Modeled	12-km grid cells
Fires (prescribed and wild)	Point	Not Modeled	12-kmgrid cells, vertical based on plume calculations

Three different approaches for spatial allocation were used to produce emission inputs for HEM-3: point, shape, and tract, where fugitives, shapes, and tracts were treated as HEM-3 "area" sources and point sources were run as HEM-3 point sources. Point sources were run at a specific latitude and longitude. Area sources were run as a flux over a specific polygon. For fugitives, the polygons are generally small (i.e., smaller than facility dimensions) and well-defined rectangular polygons based on the NEI point-source inventory and were be modeled at the specific geographic location and dimensions provided by the release-point parameters in the NEI. Shape-level emissions were new to NATA, introduced to the NEI in 2008. Shapes provide sub-county-resolution emissions data estimated by S/L/T or EPA, accounting for the fact that the emissions are limited to particular areas of the county. These were available for CMVs at ports and underway and for locomotives. For NATA, the port and underway shapes were simplified into polygons that can be input into HEM-3. The locomotive shapes were not used due to the large number of them; instead, they were allocated to tracts like the other county-level emissions.

For onroad, nonroad, and nonpoint data categories other than locomotives and CMVs, the NEI reported emissions at the county level. The NEI reported locomotives and CMVs as shapes. In the 2011 NATA, HEM-3 was run at the tract level for locomotives and sources reported at the county level, and HEM-3 was run at the NEI shapes for CMVs. Census tracts are land areas defined by the U.S. Census Bureau that vary in size and typically contain about 4,000 residents each. Census tracts are typically smaller than 2 square miles in cities, but are much larger in rural areas. Similar to previous NATAs, we spatially allocated the county-level emissions reported in NEI to the census tracts within the county. To make the county-to-tract emissions assignments, we used spatial-allocation factors derived from the distributions of various "spatial surrogates" that have geographic patterns expected to be similar to the geographic patterns of the source of the emissions. The spatial-surrogate underlying data were the same as those used for CMAQ except that (see Section 2.2), for areas outside the CMAQ domain (i.e., Alaska, Hawaii, Puerto Rico and the U.S. Virgin Islands), some of the surrogates were unavailable. In such cases, a secondary or tertiary surrogate was used.

The workbook of surrogate documentation provided in the SupplementalData folder includes the geographic extent of the spatial surrogates. We added three county-specific county-to-tract surrogates (Maricopa truck stops for extended idling, Maricopa golf courses, and Hawaii commercial lawn) and made some changes to spatial-allocation assignments as a result of the S/L/T agency comments received during the preview (which were not utilized for the CMAQ modeling).

The following data related to the census-tract surrogates can be found in the SupplementalData folder:

- Cross-references of spatial surrogates by SCC ("SCC__spatialsurrogates_xref.xlsx")
- Definitions and sources of spatial-surrogate data ("US_SpatialSurrogate_Workbook_v072115.xlsx")
- Database of census-tract spatial-surrogate fractions ("SpatialSurrogate_CensusTract_FRACTIONS_finalNATA.zip")
- County-specific census-tract fractions for commercial lawn equipment and garden equipment for Maricopa County, AZ and Honolulu, HI, and for truck stops for Maricopa County ("countyspecific_censustractsurrogates.zip").

Spatial-allocation factors for county-to-tract were generated by the Spatial Surrogate Tool, which was the same tool used to generate surrogates for county-to-grid-cells. This tool had to be updated to produce surrogates based on census tracts for NATA.

2.3.1.2 Temporal Allocation

Exhibit 21 contains descriptions of the differences between HEM-3 and CMAQ in the temporal variation of emissions used for NATA. While CMAQ read in hourly emissions for each pollutant within a 12-km grid cell, HEM-3 used different temporal keywords along with scalars that provided the variation in emissions at seasonal, monthly, day-of-week, or hourly time steps. That does not mean the temporal approach is more refined in CMAQ, as hourly emissions were not available for all sources in the inventory and temporal profiles were used to produce the hourly data. In particular, the level of detail for temporalization for point sources was the same in HEM-3 and CMAQ. For other categories, CMAQ had a more refined approach because HEM-3 was not temporalized by pollutant (other than onroad, which retained season seasonal variations at the pollutant level) and because all sources within a HEM run group (see Section 2.3.2) used the same temporal profiles (with the exception of onroad sources, which use county-specific, but not SCC-specific, temporalization across pollutants and sources).

Exhibit 21. Temporal-allocation Approach—HEM-3 versus CMAQ

Category	Resolution Available for NEI Modeling [a]	Temporal Approach for HEM-3	Temporal Approach for CMAQ
Point (including airports)	Annual	Monthly/day-of-week/hourly profiles applied to sources consistent with NATA CMAQ platform	Hourly by pollutant
Point—EGU sources	Hourly emissions for NO_X, SO_2 and hourly heat input [b]	Hourly (HOUREMIS) by facility and unit using emissions from CMAQ	Hourly by pollutant
RWC	Annual	Hourly (HOUREMIS) based on CMAQ hourly emissions (summed across SCCs) of VOCs and $PM_{2.5}$	Hourly by pollutant
Nonpoint and locomotives	Annual	Uniform monthly/day-of-week/hourly profiles based on the most-used profile across these SCCs from CMAQ platform (HROFDY)	Hourly by pollutant
CMV	Annual	Uniform monthly/day-of-week/hourly profiles based on monthly pattern from CMAQ platform (MONTH)	Hourly by pollutant
Nonroad	Monthly	Pollutant-specific seasonal variation. Uniform seasonal/monthly/day-of-week profiles. Hourly profile based on the most-used profile across these SCCs from the CMAQ platform.	Hourly by pollutant

Category	Resolution Available for NEI Modeling [a]	Temporal Approach for HEM-3	Temporal Approach for CMAQ
Onroad	Hourly	Pollutant-specific seasonal variation. Monthly/weekday-Saturday-Sunday/hourly profiles based on single pollutant	Hourly by pollutant
Agricultural burning	County-monthly	Not Modeled	Hourly by pollutant
Biogenic emissions	Hourly	Not Modeled	Hourly by pollutant
Fires (prescribed and wild)	Daily	Not Modeled	Hourly by pollutant

[a] NEI has annual resolution but some of the data that was used to create the NEI (or was available prior to inputting the data to SMOKE or the process for HEM-3) was sub-annual.

[b] From continuous emissions-monitoring data reported to EPA's Clean Air Markets Division (EPA 2015h).

2.3.2 HEM Run Groups

Exhibit 22 contains descriptions of the HEM run groups. HEM run groups provided a way to organize the sources that shared similar approaches to emission characterization, such as the release characteristics, spatial allocations, and temporal allocations. HEM run groups were further divided into source groups for purposes of providing risk results. For all but the point HEM run groups, all sources within the same HEM run group were modeled with the same release characteristics and spatial and temporal profiles. For the point HEM run groups, sources can have different temporal profiles for different facilities or sources within the facility.

Exhibit 22. HEM Run Groups Based on the Nonpoint and Nonroad NEI Data Categories

HEM Run Group	NEI Category and NATA CMAQ Platform Modeling Sector	HEM-3 Modeling Features: Release Height (RH; meters), Initial Vertical Dispersion (σ_z; meters), and Spatial and Temporal Approach	Description of Sources
NP—10meters	**NEI:** nonpoint **Platform:** some of nonpt, np_oilgas	**RH**=10 σ_z =4.7 **Spatial:** tract **Temporal:** uniform monthly/day-of-week/hourly profile 26 (smooth curve with midday peak – see Exhibit 28).	Sources: industrial processes (e.g., chemical plants, oil and gas, refineries, mines, metals); solvents (industrial surface coating;, graphic arts); fuel combustion ICI; bulk gas terminals; waste disposal
RWC	**NEI:** nonpoint **Platform:** rwc	**RH** = 6.4 σ_z = 3.2 **Spatial:** tract **Temporal:** hourly by county: sum $PM_{2.5}$ and VOC by hour and county based on the 2011 NEI v6.1 platform (the 2011 NEI v1 emissions)	Sources: fireplaces, woodstoves, hydronic heaters used for residential heating
NP—OtherLow	**NEI:** nonpoint **Platform:** part of c1c2rail, part of nonpot	**RH** = 3.9 σ_z = 3.6 **Spatial:** tract **Temporal:** same as NP—10meters	Solvents (consumer, commercial); nonindustrial surface coating (architectural coating); dry cleaning; solvent degreasing, commercial cooking; gas stations (stage 1); miscellaneous non-industrial NEC (portable gas cans, auto repair shops, structure fires, and nonpoint mercury categories such as human cremation, dental amalgam, residential fuel combustion except wood, locomotives). Note locomotive release height based on web searches.
Nonroad	**NEI:** nonroad **Platform:** nonroad	**RH** = 2 σ_z = 1 **Spatial:** tract **Temporal:** seasonal, pollutant specific. Monthly emissions were uniform within each season. Diurnal use profile 26 (it was determined to apply to 77% of risk- weighted HAP emissions and 93% of diesel PM)	Nonroad equipment such as lawn mowers, turf equipment, construction equipment, commercial generators, power-washing equipment, pleasure craft

HEM Run Group	NEI Category and NATA CMAQ Platform Modeling Sector	HEM-3 Modeling Features: Release Height (RH; meters), Initial Vertical Dispersion (σ_z; meters), and Spatial and Temporal Approach	Description of Sources
CMV with port and underway shapes	**NEI:** nonpoint **Platform:** part of c1c2rail, c3marine	**RH** = 8 (based on range of 5–10 m for c1/c2 and was chosen with input from OTAQ) σ_z = 7.4 **Spatial:** port and underway shapes in the NEI **Temporal:** Uniform monthly/day-of-week/hourly profiles based on monthly pattern from CMAQ platform (MONTH)	c1/c2 and c3 marine vessels, excluding: rail (locomotives) which is in MP-other low, and emissions in Federal Waters (FIPS=85)
Onroad Light Duty	**NEI:** onroad **Platform:** onroad, onroad_ca	**RH** = 1.3 σ_z = 1.2 **Spatial:** tract **Temporal:** seasonal temporal variation is pollutant-specific and county-specific. County-specific seasonal/monthly/day-of-week/hourly profiles were the same across all pollutants and were based on benzene hourly emissions from SMOKE-MOVES.	Passenger cars, motorcycles, light-duty trucks, and refueling of all sources (light- and heavy-duty vehicles)
Onroad Heavy Duty	**NEI:** onroad **Platform:** onroad, onroad_ca	**RH** = 3.4 σ_z = 3.2 **Spatial:** tract **Temporal:** same as above but the profiles were based on hourly emissions of $PM_{2.5}$	Heavy-duty diesel, heavy-duty gas, buses
Point, excluding airports	**NEI:** point **Platform:** part of ptnonipm, ,pt_oilgas, ptegu	**RH** and σ_z based on stack parameters for point sources. For area sources, RH based on release height and σ_z was 0. **Spatial:** point = coordinates; area = coordinates + dimensions **Temporal:** For ptnonimpm and pt_oilgas: monthly, day-of-week, diurnal based on SCC (same approach was used in 2011 modeling platform) For ptegu: hourly profiles based on sum of PM, CO, and VOC emissions	Facilities in the point inventory—for all facilities except those with facility source code = 100, emissions used at facility-unit-process-release point level. Also excluding asphalt plants with no geographic coordinates (those that move around) and facilities with no HAP emissions.

HEM Run Group	NEI Category and NATA CMAQ Platform Modeling Sector	HEM-3 Modeling Features: Release Height (RH; meters), Initial Vertical Dispersion (σ_z; meters), and Spatial and Temporal Approach	Description of Sources
Airports	**NEI:** point **Platform:** part of ptnonipm	**RH** = 3 σ_z = 3 **Spatial:** For runway-area (line) sources: length based on NTAD or OTAQ runway endpoint coordinates; 50-m width for the major airports, 25-m width for the OTAQ-provided (smaller) airports. All facility emissions (NEI) spread equally over the runway(s). For non-runway sources: 10-m-square area centered on NEI coordinates. **Temporal:** same as used in CMAQ—monthly/day-of-week/hourly profiles. Monthly and hourly profiles different for commercial versus general-aviation airports; monthly also different for Alaska seaplanes.	Facilities in the point inventory—all emissions where facility source type code = 100 (airports). Emissions used at the facility unit's process release point.

2.3.3 Point Excluding Airports

Point sources in the ptnonipm sector of the NATA platform were modeled differently from those in the ptegu sector with respect to the temporalization of the emissions. The ptegu sectors were temporalized allowing for hourly variation at the unit level, whereas the ptnonipm (including ptoil_gas) used the monthly, day-of-week, and diurnal profiles used by SMOKE. Within the run, each ptegu unit was temporalized using hourly emission values as discussed in Section 2.3.3.2. Non-EGU units were modeled in HEM-3 using temporal-allocation factors derived from the temporal profiles used in SMOKE for CMAQ. Many facilities included a mixture of EGU and non-EGU processes. In such cases, all sources at a given facility were modeled in the same HEM-3 run.5 This ensured that ambient impacts were calculated for a consistent set of receptor locations for all sources at the facility.

2.3.3.1 Point: Non-EGU

Sources in the ptnonipm sector were modeled as point sources or area sources in HEM-3, with temporal variations addressed using the variable-emission (EMISFACT) feature of HEM-3. This allows the model to read emission scalars, which adjust the annual-average emissions based on month, day-of-week, and hour-of-day. Different options were used, depending on how the emissions vary (hour-of-day only, hour-of-day and day-of-week, etc.). The actual profiles applied were the same as those applied for the SMOKE/CMAQ processing (see bullet 9d below for more details).

The following processing steps were used:

1. Temporal profiles to be used for each emission point were read from the SMOKE ptnonipm temporal-profiles report. Key data fields were the facility ID, unit ID, release point ID, process ID, monthly temporal profile code, daily temporal profile code, and hourly temporal profile code.

---

[5] AERMOD temporalization is performed at the level of source IDs, so using different temporalization schemes at one facility is possible.

2. The annual-emission file (SMOKE FF, ptnonipm) was read. Key fields were the facility ID, unit ID, release point ID, process ID, the emission release type code, SCC, pollutant code, annual emissions, latitude, longitude, stack release point parameters, and fugitive release point emission parameters.

3. Annual emission file records with a facility source code of "100" (airports) were excluded. Airport emission records were processed separately.

4. The annual-emission file is matched to the temporal file using a concatenation of facility ID, unit ID, release point ID, and process ID. The monthly, daily, and hourly temporal profile codes were then copied into the annual file.

5. The annual file was matched to a pollutant cross-reference file, which links the pollutant codes (primarily CAS numbers) used in the NEI to the pollutant library used in HEM-3. The cross-reference file, developed for the 2005 NATA and Risk and Technology Review (RTR; EPA 2015i) modeling efforts, was updated for the 2011 NATA relying on the valid codes in the pollutant-code table in the EIS. HEM-3 HAP identifiers were copied to the annual-emissions file, along with factors that will be used to adjust emissions of certain HAP compounds to reflect the fraction of HAP content. For instance, chromic acid (H_2CrO_4) emissions were multiplied by a factor of 0.4406 so that only the chromium portion of the molecule mass will be compared with the dose-response value. In this step, different NEI HAP species may be linked to the same HEM-3 pollutant, so that emissions will be combined in the modeling step. For example: glycol ether species were put in the lumped pollutant category called "glycol ethers," specific polychlorinated biphenyls were put into the category "Polychlorinated biphenyls," and hydrogen cyanide and cyanide were lumped into "cyanides."

6. Missing source parameters were defaulted for fugitive sources (emission release type code = 01). Length and width values less than one meter were also replaced. The following defaults were applied:

    a) If width or length was missing, the parameter was assigned a value of 10 m.

    b) If width or length was less than 1 m, the parameter was set to 1 m.

    c) If width, length, and height were all missing, height was set to 3.05 m, and length and width were each set to 10 m.

    d) If height was missing but length and width were populated with nonzero values, height was left at 0.

    e) If angle was missing, it was left at 0.

7. The NEI had no missing parameters for vertical stacks (i.e., emission release type code = 02). For emission release types of 03-06 (i.e., non-fugitive releases), there may be missing parameters. In those cases, missing vertical stack parameters were defaulted as follows:

    a) Exit-gas velocity: computed from flow rate and diameter if those parameters were available, otherwise used SCC-specific[6] or global SMOKE default (4 m/s) value.

    b) Exit-gas temperature: used SCC-specific or global SMOKE default 295.4 K) value.

---

[6] The SCC defaults, which are in metric units (i.e., m, m/s, and K) are in the "PSTK" ancillary file (pstk_20nov2006_v0.txt) which is available from the 2011v6.2 platform ancillary data (specifically here).

   c) Inside release-point diameter: used SCC-specific or global SMOKE default (0.2 m) value.

   d) Release-point height: used SCC-specific or global SMOKE default (3 m) value.

8. Within each facility, an alphanumeric source ID that was no more than eight characters in length was assigned to each unique combination of unit ID, process ID, and emission release point ID. The source ID sequence began with S0000001 for each facility. The linkage of source ID to inventory IDs (i.e., unit ID, release Point ID, process ID, and emission release type code) was archived. (Source IDs were unique within a particular facility, but not among facilities.)

9. An intermediate file was written with a separate record for each source ID and for each HEM-3 pollutant (consolidating some inventory pollutants such as glycol ethers). This was used to create four separate files used by HEM-3, as follows:

   a) Facility list options file, one record per facility. This file contained the facility ID, along with model settings to be used in HEM-3, such as the overall radius of the modeling domain, the cutoff distance for discrete-receptor modeling, and the specifications of the polar grid to be used for interpolation calculations in the region beyond the discrete modeling cutoff. Exhibit 23 contains a description of the fields in the facility list options file.

Exhibit 23. Fields in the HEM-3 Facility List Options File

Field	Default Setting (if field left blank)	Description
Facility ID		Alphanumeric string identifying the facility being modeled; up to 30 characters long.
Met station	Met station selected by model as closest to the facility	The name of the meteorological surface station to be used by HEM-3 when modeling each facility; up to 20 characters long; generally chosen by model but you have the option of specifying.
Rural/Urban	D for default	Used to set the type of dispersion environment for HEM-3. "R" indicates rural land use surrounding the facility; "U" indicates urban land use; and "D" indicates the default setting under which the model will find the nearest Census block to the facility center and determine whether that Census block is located in an urbanized area as designated by the 2010 Census. [Note: The default setting when using the 2000 Census is always rural.]
Max distance	50,000 m	The outside maximum radius of the modeling domain (≤ 50 km).
Modeling distance	3,000 m	The cutoff distance for individual modeling of ambient impacts at census blocks; beyond this distance, ambient impacts were interpolated rather than explicitly modeled. [Note: For polygon source types, set the modeling distance > the largest distance across the polygon.]
Radials	16	The number of radials in the polar-receptor network emanating from the facility center.
Circles	13	The number of concentric circles in the polar receptor network, centered on the facility center.
Overlap distance	30 m	The distance, measured from each emission source at a facility, at and below which a source and receptor are considered to be overlapping. Must be ≤ 500 m.

Field	Default Setting (if field left blank)	Description
Acute	N	Selecting "Y" directs the model to include short-term (acute) concentration calculations and hazard predictions.
Hours	1 hour	The short-term (acute) averaging period that HEM-3 should use for ambient concentrations. Four averaging period options are available: 1, 6, 8, and 26 hours.
Elevations	Y	Elevations of receptors were accounted for by default; selecting "N" excludes elevations from the model run.
Multiplier	10	The acute multiplier applied to the average emission rate and used to approximate the short-term emission rate (e.g., 10 times the rate entered in the HAP Emissions input file). Multi-facility HEM-3 assumed that this short-term rate could occur at the same time as the worst-case meteorological conditions, making the acute results conservative estimates.
First ring distance (ring1)	Calculated by model to be just outside the source locations, but not less than 100 m from facility center	The distance to the first ring (circle) of the polar network as measured from the facility center. You can override the default distance calculated by multi-facility HEM-3 to fit the size and shape of the facility properties to be modeled.
Deposition (dep)	N	Deposition was not modeled by default. Selecting "Y" directs the model to calculate deposition in the model run (particle, vapor, or both as designated below). Note: if not modeling deposition, ignore the depletion, phase, and particle and vapor deposition fields (below).
Depletion (depl)	[depletion modeled automatically with deposition]	The current model automatically depletes the calculated deposition flux from the ambient concentrations if you opt to calculate deposition; therefore no entry necessary.
Phase	B	The default value "B" directs the model that both particles and vapor deposition will be modeled; use "P" for particle-only deposition modeling; use "V" for vapor-only deposition modeling. Value must be consistent with emissions (e.g., do not use "B" if emissions are 100% "P" or "V").
Particle Deposition (pdep)	WD for wet and dry particle deposition	The default value "WD" directs the model to incorporate both wet and dry deposition for particles. Use "WO" for wet only; use "DO" for dry only; use "NO" if not modeling deposition of particles.
Particle Depletion (pdepl)	[particle depletion modeled automatically with particle deposition]	The current model automatically depletes the calculated deposition flux for particles from the ambient concentrations if you opt to calculate particle deposition; therefore, you need not enter anything in this column.
Vapor Deposition (vdep)	WD	The default value "'WD" directs the model to incorporate both wet and dry deposition for vapor pollutants; use "WO" for wet only; use "DO" for dry only; use "NO" if not modeling deposition of vapor pollutants.
Vapor Depletion (vdepl)	[vapor depletion modeled automatically with vapor deposition]	The current model automatically depletes the calculated deposition flux for vapor pollutants from the ambient concentrations if you opt to calculate vapor deposition; therefore, you need not enter anything in this column.
All Receptors	Y	"Y" directs model to calculate results for all receptors by pollutant and source. Select "N" to receive pollutant and source contributions for the maximum-populated and maximum off-site receptors only.
User receptors	N	Select "Y" to include user receptors in a separate input file.

Field	Default Setting (if field left blank)	Description
Building Downwash (bldg_dw)	N	Selecting "Y" directs the model to include building-downwash calculations in the model run. Note: if you are modeling building downwash, building-dimension information is required in a separate input file.
Urban Population	None; only needed if "U" specified in Rural/Urban field	If you indicate "U" for urban land use (in Rural/Urban field above), then you must provide model with the urban population size, otherwise leave blank. Note: if you specify "U" in the Rural/Urban field but provide no urban population value in this field, the model will re-set your "U" to default.
FASTALL	N	FASTALL was not used by default. Selecting "Y" directs the model to use the control option FASTALL, which conserves model run time by simplifying dispersion algorithms.

b) Stack parameter file, with one record for each facility and source ID (as defined in step 8). This file specifies the source latitude and longitude, the source type (A = area, P = point), and stack parameters for point sources or fugitive-release parameters for area sources. Source types were assigned based on the emission release point type code in the inventory, as shown in Exhibit 24. Note that the stack- and area-source parameters were converted from English units (used in the NEI) to metric units (used in HEM-3; e.g., meters, meters per second, and Kelvin rather than feet, feet per second, and Fahrenheit).

Exhibit 24. HEM-3 Assignments of Emission Release Point Type

Emission Release Point Type Code	Emission Release Point Type Description	HEM-3 Source Type code
01	Fugitive	A
02	Vertical	P
03	Horizontal	P
04	Goose Neck	P
05	Vertical with Rain Cap	P
06	Downward-facing vent	P
99	Unknown	P

c) Pollutant emissions file, with one record per combination of facility ID, source ID, and pollutant, giving annual emissions in TPY.

d) Temporal allocation file, with records for each facility ID and source ID. These records give scalar factors that should be applied to the annual-average emission rate to compute hourly emissions. (The scalar factors are dimensionless.) HEM-3 has a number of options for the temporal-allocation records, depending on how the emissions vary. We checked the SMOKE temporal profiles to see which HEM-3 keyword should be used. The following keyword options were used for ptnonipm depending upon value of profile codes:

- MHRDOW: emissions vary by month, day type, and hour. Scalar = (monthly factor/average monthly factor) × (daily factor/average daily factor) × (hourly factor/average hourly factor)
- HRDOW: emissions vary by day type and hour (monthly profile code 262, uniform through the year). Scalar = (daily factor/average daily factor) × (hourly factor/average hourly factor)
- HROFDAY: emissions vary by hour only (monthly profile code 262, daily profile code 7, both uniform). Scalar = (hourly factor/average hourly factor)
- MONTH: emissions vary by month only (daily profile code 7, hourly profile code 24, uniform). Scalar = (monthly factor/average monthly factor)

2.3.3.2 Point: EGUs

EGUs were treated similarly to the ptnonipm sector other than for temporalization (see bullet 9d from the previous section). Ptegu sources were modeled as point sources or area sources (fugitives) in HEM-3, with temporal variations addressed using the hourly-emission (HOUREMIS) feature of HEM-3. The same processing steps as for ptnonipm were followed, up until step 9d. For temporalization, the following steps were applied:

1. Hourly-emissions data for CAPs were read from 365 SMOKE hourly-emissions reports, one per day. Each daily file contained the hourly-emission values for that day for all CAPs. This covered a large regional data set, but only the records matching the batch filter (from step 1) were read. The following fields were read: date, hour, data source ID, region code (FIPS), state, county, SCC, SIC, facility ID, unit ID (char 1), release point ID (char 2), process ID (char 3), plant name, CO tons/hour, NO_X tons/hour, SO_2 tons/hour, $PM_{2.5}$ tons/hour, and VOC tons/hour.
2. Hours in the SMOKE hourly-emission reports were in Greenwich Mean Time (GMT). The hours were shifted from GMT to local time using the time zone of the county.
3. A separate facility temporal file was created for each source location and stack height covered in the hourly-emissions data files, and data from all 365 daily files were copied to this file. Sources with the same location and stack parameters were lumped together for this algorithm. For each hour of the year, an hourly emission scalar value was computed, as follows:

$$Scalar_{hour\ i} = \frac{(VOC + CO + PM_{2.5})_{hour\ i}}{\sum_{all\ hours}(VOC + CO + PM_{2.5})}$$

4. Source IDs for each facility were merged from the annual inventory.
5. For each facility, the source IDs and scalar factors were written to a text file suitable for the HEM-3 HOUREMIS option.

2.3.3.3 Special Considerations for Coke Ovens

For the 2011 NATA, EPA used the same approach developed for the 2005 NATA to account for enhanced buoyancy for coke oven emissions. Enhanced buoyancy around hot banks causes the associated emission plume to travel a significant vertical distance before it disperses laterally. This buoyancy was accounted for in the NATA modeling by increasing the modeled release height of certain emission stacks. Any coke oven stacks associated with charge lids, doors, and charging, pushing, or off-take processes, and with stacks shorter than 126 ft, were set to 126 ft. This height was chosen because most (160 of 165) of the coke oven stacks modeled in 2002 and earlier NATAs had stacks shorter than 126 ft (heights overall ranged from 10 ft to 315 ft).

2.3.4 Point: Airports

Airports were inventoried in the NEI as point sources. Their emissions were part of the ptnonipm sector of the NATA modeling platform and were provided with other sources in that sector. They were modeled in HEM-3 as line sources where runway endpoint data were available, and as small area sources in all other cases. Where line sources were modeled, the line-source option in HEM-3 was used, which creates an area source using line endpoints, source width, and other source parameters. In all cases, temporal variations were addressed using the variable-emission (EMISFACT) feature of HEM-3. This allowed the model to read emissions scalars, which adjust the annual-average emissions based on month, day-of-week, and hour-of-day. Airport-emission records were extracted from the annual ptnonipm emission file based on the facility source code. In the NEI, a facility source code of "100" indicated an airport. All sources of emissions contained within an airport facility (i.e., aircraft, ground-support equipment, auxiliary-power units, and any other source reported by S/L/T agencies at an airport facility) were summed and modeled the same way.

For airports modeled as runway line sources, the NEI geographic coordinates were not used directly in the modeling of the airport emissions, but instead they were used to ensure the runway location data were assigned correctly to EIS facilities. The runway-location data were from two sources. One source was the public-use-airports dataset from the 2012 National Transportation Atlas Database (NTAD), developed by the Bureau of Transportation Statistics of the U.S. Department of Transportation.[7] The NTAD is a set of nationwide geographic databases of transportation facilities, transportation networks, and associated infrastructure. The NTAD shapefile for public-use airports was derived using textual (non-spatial) data from the Federal Aviation Administration (FAA) National Airspace System Resource Aeronautical Data. The NTAD airport dataset contained data on approximately 6,600 runways at 4,600 airports. Using GIS software, EPA's Office of Air Quality Planning and Standards (OAQPS) converted these lines to endpoints, which were the source location inputs for the line-source option in HEM-3. The EIS IDs were assigned to these airports by comparing identification data in the EIS to airport-facility data in FAA Form 5010 records (FAA 2015) and other FAA databases, the FAA's Emissions and Dispersion Modeling System, state-supplied data, and online tools (e.g., Great Circle Search[8]). Where there was no match, we used location coordinates to determine if airports were geographically close enough to be the same facility. The typical runway width in the NTAD data was 50 m, and this value was used as the width for all runways for the modeling of the NTAD runways.

The second source of runway location data used to derive line-source inputs was a dataset developed by OTAQ. For airports not included in the NTAD dataset, OTAQ used the same FAA textual data that were used to derive the NTAD data to create runway endpoints for the smaller airports that constitute the OTAQ dataset. The OTAQ airport dataset contained data on approximately 9,000 airports, and in almost all cases, there was one runway per airport. For the smaller airports, the FAA data typically included two runway endpoints, a single endpoint with a runway length and heading, or a single location (presumed to be the runway center) with a runway length and heading. These data were used by OTAQ to derive runway endpoints, which were used as the source-location inputs for the line-source option in HEM-3. The EIS IDs were assigned to these airports in the same way as the airports in the NTAD dataset. The widths of the runways in the OTAQ dataset were smaller than those in the NTAD data, and a width of 25 m was used for all runways for the modeling of the OTAQ runways.

---

[7] See this site.

[8] See this site.

Airports without runways (e.g., heliports) or without sufficient data to generate runway endpoints were modeled as small area sources (10-m square) with the NEI location coordinates as the southwest corner. During the NATA review, it was discovered that the seaplane runways were often over land, so seaplanes with runway endpoints were all re-modeled as small area sources. While the small area source may characterize a helipad fairly well, it likely would not for a runway. However, without good data on runways it was difficult to characterize such sources. Modeling a larger area source (or line source) could result in the source overlapping nearby population receptors, resulting in overestimates of exposure. Further, the emissions for the smaller airports are likely much lower. Therefore, the small area source was used for all airports not included in the NTAD or OTAQ datasets. There were some airports in the NTAD and OTAQ datasets with runway geographic coordinates significantly different from the coordinates in the NEI for the same airports. Coordinate differences of 2-to-3 km were expected because airports are generally large and the runway coordinates may be far from the single coordinate set that represented the airport in the NEI. There were 112 airports in the NTAD and OTAQ datasets with coordinates more than 2,500 m away from the NEI coordinates, and we modeled these airports the same way we modeled airports without runways or without sufficient data to generate runway endpoints.

The file "Airport List.xlsx", provided in the SupplementalData folder, includes all the airports modeled and indicates which airports were modeled using runway endpoints (NTAD and OTAQ) and which were modeled as area sources (EIS). The list of seaplanes is in "Airport List_seaplaneispoint.xlsx."

For some airports, census block receptors fell on or near the runway or within or near the small area source. However, people do not live within the airport boundaries or close to runways. Thus, blocks that fell either within 30 m of the geographic coordinates of the area source or within 50 m of the runway are assigned the next-highest receptor concentration of a receptor that lies outside these distances. This approach was the same taken for other point sources and ports.

2.3.4.1 Temporal Profiles Assigned to Airports

Special care was taken to assign realistic operating conditions to airports. For Alaska seaplanes, it was found (Meredith Pedde, personal communication) that they operate daily other than in winter (November thru March), when they run two days per week. Thus, we assigned the monthly profile in Exhibit 25 to all airport facilities in Alaska with a description of "seaplane." A database of EIS facilities with facility descriptions, "EISAirportwDescription.accdb, is provided in the SupplementalData folder.

Exhibit 25. Monthly Temporal Profile for Alaska Seaplanes (Counts and Percentages)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Sum
20 3.4%	20 3.4%	20 3.4%	70 11.9%	70 11.9%	70 11.9%	70 11.9%	70 11.9%	70 11.9%	70 11.9%	20 3.4%	20 3.4%	590 100%

We also looked at the diurnal profiles. Meredith suggested that for general aviation we use a diurnal profile that has no activity between 10 p.m. and 6 a.m. (due no tower operation). We used this profile for airports that were: 1) not commercial (based on list from Meredith Pedde) and 2) not heliports. A complete list is in "general aviation airports for 6amto10pm profile.xlsx" in the SupplementalData folder. It was developed by taking all airports from the FF10 (facility type = 100), merging in the list of commercial airports provided by OTAQ (Meredith Pedde), removing them, and removing seaplanes (using the facility description file, EISAirportwDescription.accdb, discussed above).

This new profile (shown in Exhibit 26) was developed by using the diurnal profile weights (1st row in below table), setting hours to 0 between 10 p.m. and 6 a.m. (2nd row), and then renormalizing (3rd row)

Exhibit 26. Diurnal Temporal Profile for General Aviation (Counts, Zero-outs, and Final Percentages)

Hr 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
198	186	182	187	210	250	311	388	467	528	571	604	620	631	635	624	594	548	531	509	425	327	257	218
0	0	0	0	0	0	311	388	467	528	571	607	650	631	635	624	594	548	531	509	425	327	0	0
0%	0%	0%	0%	0%	0%	3.741%	4.667%	5.618%	6.351%	6.869%	7.266%	7.458%	7.591%	7.639%	7.506%	7.145%	6.592%	6.388%	6.123%	5.112%	3.934%	0%	0%

2.3.4.2 Lead Adjustment for Piston Aircraft

The NEI estimated that lead emissions from piston aircraft totaled 486 TPY, nationally. Of these emissions, the NEI provides estimates of lead emitted during several modes of operation at and near airports, which total 248 TPY. These estimates were used in NATA. The NEI also estimated in-flight lead emissions at the state level (238 TPY) which was not included in NATA. Lead emissions near airports included lead emitted during the climb-out and approach modes, which occur at altitudes and are not included in NATA. To account for this, we adjusted down by 50 percent the NEI-specific emissions estimates used in the HEM-3 modeling, based on previous modeling conducted at the Santa Monica (SMO) airport indicating that nearly 50 percent of emissions occurred in these higher-altitude modes (see Exhibit 27).

Exhibit 27. Lead Emissions (kg/yr) at SMO in 2008, by Aircraft Operation Mode

Mode	Emissions (% of Total)
Taxi To Runway	20.4 (17.6%)
Run-up	13.5 (11.4%)
Takeoff Roll	10.0 (8.4%)
Climb-out	37.9 (32.7%)
Approach	17.9 (15.8%)
Landing	9.4 (7.9%)
Taxi to Apron	9.5 (8.4%)

The processing of data for each type of airport source for HEM-3 is described below.

Line Source Airports

1. Within each airport facility, an alphanumeric source ID that was no more than eight characters in length was assigned to each unique combination of category name (an identifier of the source type used in HEM-3, e.g., APL for airport line) and runway ID. For each facility, the source ID sequence began with APL00001. Source IDs were unique within a particular facility, but not among facilities.

2. For each airport facility, airport emission records were linked to the airports in the runway data files by the facility ID, and the annual emissions were apportioned to each runway-by-runway area.

3. Temporal profiles to be used for each emissions source were read from the SMOKE ptnonipm temporal-profiles report. Key data fields were the facility ID, unit ID, release point ID, process ID, monthly temporal profile code, daily temporal profile code, and hourly temporal profile code.

4. Key fields from the annual airport file were the facility ID, runway ID, unit ID, release point ID, process ID, lat/long of the beginning of the runway, lat/long of the end of the runway, pollutant code, and annual emissions.

5. The annual airport emission file was matched to the temporal file using a concatenation of facility ID, unit ID, release Point ID, and process ID. The monthly, daily, and hourly temporal profile codes were then copied into the annual file.

6. The annual airport file was matched to a pollutant cross-reference file, which linked the pollutant codes (primarily CAS numbers) used in the NEI to the pollutant library used in HEM-3. The cross-reference file was developed for the 2005 NATA and RTR modeling efforts, and updated for the 2011 NATA relying on the valid codes in the pollutant-code table in the EIS. HEM-3 HAP identifiers were copied to the annual-emissions file, along with factors, which will be used to adjust emissions of certain HAP compounds to reflect the fraction of HAP content.

7. Line-source release parameters were assigned to each record as follows: release height = 3 m, σ_z = 3 m, and runway width = 50 m for NTAD runways and 25 m for OTAQ runways.

8. An intermediate file as written with a separate record for each facility ID and source ID and for each HEM-3 pollutant (consolidating some inventory pollutants such as glycol ethers). This was used to create four separate files used by HEM-3, as follows:

    a) Facility list file, with one record per facility. This file contained the facility ID, along with model settings to be used in HEM-3, such as the overall radius of the modeling domain, the cutoff distance for discrete-receptor modeling, and the specifications of the polar grid to be used for interpolation calculations in the region beyond the discrete modeling cutoff.

    b) Stack parameter file, with one record for each facility ID and source ID. This file specified the source type (N = line), beginning and ending lat/long of the line (runway), and release parameters for line sources (release height, σ_z, and width).

    c) Pollutant emissions file, with one record per combination of facility ID, source ID, and pollutant, giving annual emissions in TPY.

    d) Temporal allocation file, with one record for each facility ID and source ID. These records gave dimensionless scalar factors, which should be applied to the annual-average emission rate to compute hourly emissions. HEM-3 had a number of options for the temporal allocation records, depending on how the emissions varied. We checked the SMOKE temporal profiles to see which HEM-3 keyword should be used. The following keyword options were used for ptnonipm:

        - MHRDOW: emissions vary by month, day-type, and hour: Scalar = (monthly factor/average monthly factor) × (daily factor/average daily factor) × (hourly factor/average hourly factor)

- HRDOW: emissions vary by day-type and hour (monthly profile code 262, uniform through the year). Scalar = (daily factor/average daily factor) × (hourly factor/average hourly factor)
- HROFDAY: emissions vary by hour only (monthly profile code 262, daily profile code 7, both uniform). Scalar = (hourly factor/average hourly factor)
- MONTH: emissions vary by month only (daily profile code 7, hourly profile code 24, uniform). Scalar = (monthly factor/average monthly factor)

Small Area Source Airports

1. Within each airport facility, an alphanumeric source ID that was no more than eight characters in length was assigned to each unique combination of unit ID, process ID, and emission release point ID. The source ID sequence began with APP00001 for each facility. Source IDs were unique within a particular facility, but not among facilities.
2. Temporal profiles to be used for each emissions source were read from the SMOKE ptnonipm temporal-profiles report. Key data fields were the facility ID, unit ID, release point ID, process ID, monthly temporal profile code, daily temporal profile code, and hourly temporal profile code.
3. Key fields from the annual airport file were the facility ID, unit ID, release point ID, process ID, pollutant code, annual emissions, latitude, and longitude.
4. The annual-airport-emission file was matched to the temporal file using a concatenation of facility ID, unit ID, release point ID, and process ID. The monthly, daily, and hourly temporal-profile codes were then copied into the annual file.
5. The annual airport file was matched to a pollutant cross-reference file, which linked the pollutant codes (primarily CAS numbers) used in the NEI to the pollutant library used in HEM-3. The cross-reference file was developed for the 2005 NATA and RTR modeling efforts, and it was updated for the 2011 NATA relying on the valid codes in the pollutant code table in the EIS. HEM-3 HAP identifiers were copied to the annual-emissions file, along with factors, which will be used to adjust emissions of certain HAP compounds to reflect the fraction of HAP content.
6. An intermediate file was written with a separate record for each facility ID and source ID and for each HEM-3 pollutant (consolidating some inventory pollutants such as glycol ethers). This was used to create four separate files used by HEM-3, as follows:
    a) Facility list file, with one record per facility. This file contained the facility ID, along with model settings to be used in HEM-3, such as the overall radius of the modeling domain, the cutoff distance for discrete-receptor modeling, and the specifications of the polar grid to be used for interpolation calculations in the region beyond the discrete modeling cutoff.
    b) Stack parameter file, with one record for each facility ID and source ID. This file specified the source type (A = area), lat/long of the southwest corner of the area source, and release parameters for area sources (length of the sides in the x and y directions, angle, release height, σ_z).
    c) Pollutant emissions file, with one record per combination of facility ID, source ID, and pollutant, giving annual emissions in TPY.

d) Temporal allocation file, with one record for each facility ID and source ID. These records gave dimensionless scalar factors, which should be applied to the annual-average emission rate to compute hourly emissions. HEM-3 had a number of options for the temporal-allocation records, depending on how the emissions varied. We checked the SMOKE temporal profiles to see which HEM-3 keyword should be used. The following keyword options were used for ptnonipm:

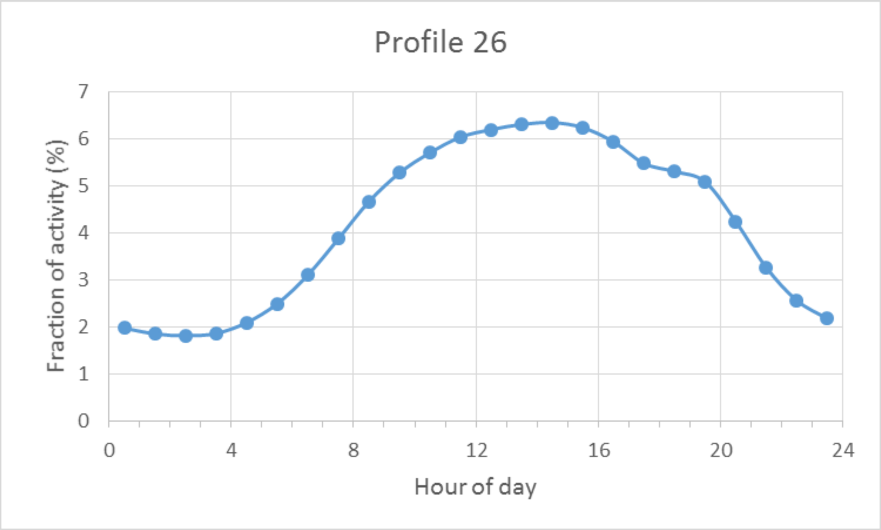
   - MHRDOW: emissions vary by month, day-type, and hour. Scalar = (monthly factor/average monthly factor) × (daily factor/average daily factor) × (hourly factor/average hourly factor)
   - HRDOW: emissions vary by day-type and hour (monthly profile code 262, uniform through the year). Scalar = (daily factor/average daily factor) × (hourly factor/average hourly factor)
   - HROFDAY: emissions vary by hour only (monthly profile code 262, daily profile code 7, both uniform). Scalar = (hourly factor/average hourly factor)
   - MONTH: emissions vary by month only (daily profile code 7, hourly profile code 24, uniform). Scalar = (monthly factor/average monthly factor)

2.3.5 Nonpoint HEM Run Groups: NP10m and NPOtherLow

The NP10m and NPOtherLow HEM run groups constituted most of the anthropogenic sources in the NEI nonpoint data category. Note that the nonroad source of locomotives was included in NPOtherLow. These were run as separate run groups due to their different release characteristics. NP10m used a 10-m release height whereas NPOtherLow used a 3.9-m release height.

The emissions were allocated to census tracts such that the sources were modeled as tract areas, with emissions based on the fraction of emissions assigned from the county to the tract based on the spatial surrogate. These two HEM run groups used the same diurnal temporal profile (SMOKE profile code 26, which was also used for nonroad sources and is shown in Exhibit 28).

Exhibit 28. Hourly Pattern of Activity for SMOKE Profile 26



2.3.6 Nonpoint HEM Run Groups: CMVs

CMVs consist of two sources groups, which are also the HEM run groups: ports and underway. Both groups were modeled with an 8-m release height, and both used a monthly temporal profile derived from SMOKE profile 19531.

The majority of the emissions from these groups were modeled at the port or underway shapes (GIS polygons) that they had been assigned to in the NEI. To reduce computational requirements for HEM-3, GIS was used to smooth and sometimes split the shapes, thereby reducing the number of vertices of each shape. For shapes that were split into multiple parts, the emissions for the shape were allocated to each part according to area.

A small amount of county-level CMV emissions was not assigned to port or underway shapes in the NEI; we assigned them to water-body shapes such as large lakes and rivers within the county. These water body shapes were modeled in the same way as the port and underway shapes.

To gap-fill missing 2011 CMV emissions in the Puerto Rico municipos San Juan, Ponce, and Guaynabo, the corresponding 2005 tract-level CMV emissions from the 2005 NATA were used. Guaynabo's 2011 emissions resulted from a change of census-tract boundaries between 2005 and 2011. Emissions from 2005 were divided up into CMV diesel, CMV residual fuel, and diesel pleasure craft, and they were already allocated to tracts from the 2000 Census. All of the 2005 CMV emissions were assigned to the ports source group, and the 2000 Census tracts were mapped to 2010 Census tracts by using GIS to determine the area percentage of the 2000 tracts in the 2010 tracts. Also, 2005 CMV POM air toxics were assigned to "PAH, total" (NEI pollutant code 130498292). The modeling results ("chai/Q") for Puerto Rico for the NP10m HEM run group were used to compute the CMV HEM-3 concentrations for 2011 for the gap-filled municipos.

The HEM-3 CMV outputs (ports and underway) were checked to see if any receptor overlapped a port or underway shape. The concentration at any overlapped receptor was assigned the concentration from the maximum non-overlapped receptor.

2.3.7 Nonpoint HEM Run Groups: RWC

RWC was both a HEM run group and source group, allowing the tract-level risks to be apportioned out to this group (but not to individual sources within the group). It was consistent with the RWC CMAQ platform sector, and it included the sources listed in Exhibit 12.

The HEM-3 release height and initial vertical dispersion parameter σ_z were computed based on a 50/50 split of 1- and 2-story houses in a tract. Release heights were set to 50 percent of the initial plume height from the chimneys, with plume being based on the sum of the chimney height plus plume rise. The vertical dispersion, σ_z, was set equal to the plume height divided by 4.3, consistent with guidance for area sources in the AERMOD User's Guide. Plume-rise estimates were based on a series of AERMOD simulations for 1- and 2-story houses with different chimney-to-house orientations (chimney on long edge of house, chimney on short edge of house, etc.) and chimney temperatures (394 and 588 K). The dimensions of the houses are:

Number of stories	Length (m)	Width (m)	Roof height (m)	Chimney height (m)
1	15.24	9.14	6.1	7.0
2	12.2	7.3	10.1	11.0

The resulting release height and σ_z for the RWC emissions were 6.4 m and 3.2 m, respectively.

This HEM run group was spatially allocated from the county to the tract using the spatial surrogates 165 and 300, described below.

Surrogate ID	Surrogate	Description	Source of Data	Vintage
165	0.5 Residential Heating–Wood, plus 0.5 Low Intensity Residential	Combination of 50% Residential Heating–Wood and 50% Low Intensity Residential	American Community Survey (U.S. Census Bureau) and National Land Cover Database	2005–2010
300	Low Intensity Residential	2006 NLCD 30-m data resampled to 250-m resolution for land-use class 22 (Developed, Low Intensity); areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20% to 49% percent of total cover. These areas most commonly include single-family housing units	National Land Cover Database	2006 Landsat Imagery

All SCCs were allocated the using the same surrogate (either 300 or 165)[9]; the national default was 165 but for some counties it was found that surrogate 165 concentrated emissions to small tracts within urban counties. For these counties, a county-specific surrogate of 300 was used. Additional urban counties were assigned the 300 surrogate to HEM-3 after CMAQ was already run.

Annual emissions were allocated using an hourly profile that was created for each county based on the sum of the hourly emissions of $PM_{2.5}$ and VOCs across all RWC SCCs, based on the approach used for the 2011v6.1 platform (processing of the 2011 NEI v1. All SCCs for the HEM-3 runs used the same profile, a weighted average profile based on the summed PM and VOC emissions.

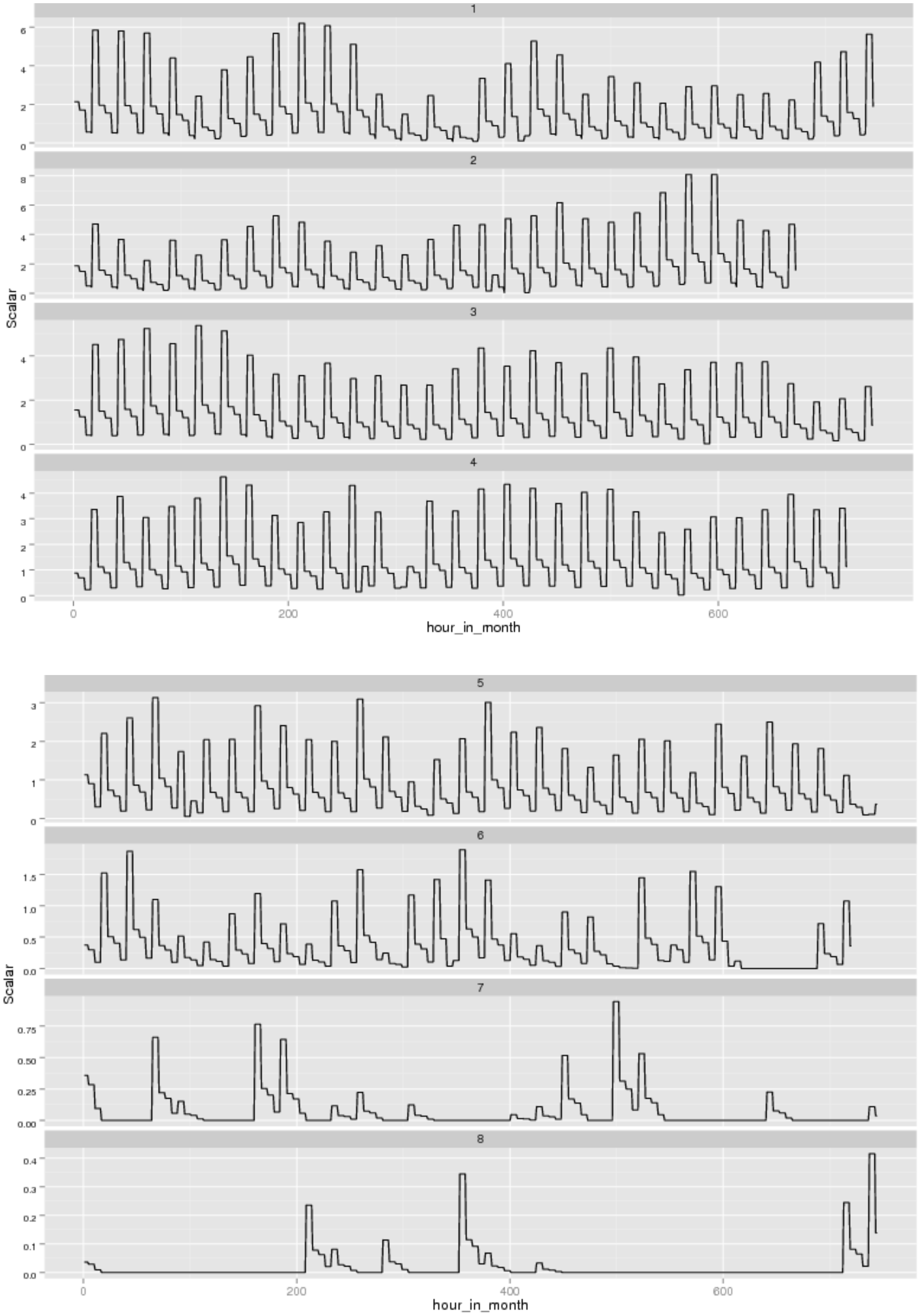
Hourly emissions were estimated from (1) an annual-to-day profile based on the daily-minimum temperature and (2) a diurnal profile putting more of the emissions in the evening and morning when people are typically using these sources. The meteorological-based approach is documented on pages 57-61 of this document.

This temporalization changed slightly between v1 and v2 in CMAQ due to the assignment of some RWC SCCs to a uniform diurnal profile and due to a fix in the SMOKE code that shifted the day-specific profile (created by SMOKE based on the daily-minimum temperature) and the hourly profile (based on the diurnal variation) to a consistent time basis (GMT). This change was not implemented for HEM-3, however. To avoid rerunning the model, we instead applied the new platform emissions to the "chai/Q" values output by HEM-3 based on the v1 temporalization. An example of this hourly, day-specific temporalization for King County, WA is provided in Exhibit 29 for January–April (months 1–4) and May–August (months 5–8).

---

[9] We inadvertently assigned fireplaces to 165, but they should have used use 300 for all counties.

Exhibit 29. Example of RWC Temporal-scaling Factors, January(1)–April(4) (top) and May(5)–August(8) (bottom), for King County, Washington



2.3.8 Nonroad HEM Run Group

Nonroad emissions were run as a single HEM run group. The sources in this group were the same as the sources in the nonroad data sector in the NEI and the nonroad platform sector. For temporalization, we used diurnal profile 26 (see Exhibit 28) along with monthly factors based on toxicity-weighted emissions from the FF10 nonroad file.

2.3.9 Onroad HEM Run Groups: Light Duty and Heavy Duty

The onroad emissions included light- and heavy-duty-vehicle emissions for on- and off-network and refueling, consistent with the onroad emissions used for CMAQ. The emissions data came from a MOVES2014 and SMOKE-MOVES run done for the NATA platform for CMAQ. As part of that run, an FF10 dataset of monthly and annual emissions was produced with all CAPs and HAPs, and sufficient SCC specificity to support spatial allocation using the same surrogates as used for the NATA platform for CMAQ. Diesel PM was taken as all-PM_{10} from the diesel-exhaust-related SCCs (see Exhibit B-3 of Appendix B).

For the HEM-3 runs, onroad emissions were temporalized separately for light- and heavy-duty vehicles to account for monthly, day-of-week, and hourly variations using the HEM-3 MHRDOW (monthly/day-of-week/hourly) option. Although seasonal variations were specific to each pollutant and profiles also varied by county, season-to-month, day-of-week, and hourly variations were the same for all pollutants within the HEM run group.

All refueling (from light or heavy) was put into light-duty vehicles because refueling is a source group and we did not want to have source groups that were in multiple HEM run groups. In addition, heavy-duty refueling was very small compared to other onroad sources. Thus, all pollutants received the same monthly, day-of-week, and hourly temporal variations. However, we split the runs into four seasons (i.e., December–February, March–May, June–August, and September–November) to allow pollutant-specific seasonal variation at the seasonal level.

MHRDOW factors were created from both the FF10 file (for the monthly factor only) and a specially created file of county/SCC/hour totals (in tons/hour) generated from the SMOKE-MOVES model (running MOVES2014) based on the NATA platform. The hourly emissions data were created for one week per month and converted to local time (taking into account daylight saving where appropriate). For each month, the period was late evening on the 21st through late evening on the 28th, local time. (In GMT, the timeframe covers the 22nd 0:00 through 28th 23:00.)

Benzene was used for light duty and $PM_{2.5}$ for heavy duty. Emissions are aggregated to the HEM run groups via a crosswalk of SCCs to HEM run group (see Appendix D or the spreadsheet file "NP_NR_OR_SourceGroup_to_SCC_CrossReference.xlsx" in the SupplementalData folder). The following calculations were made to develop the scalar inputs to HEM-3, using keyword MHRDOW to provide county-specific, monthly/day-of-week/hourly scalars.

$$Scalar = \frac{monthly\ factor}{average\ monthly\ factor} \times \frac{daily\ factor}{average\ daily\ factor} \times \frac{hourly\ factor}{average\ hourly\ factor}$$

$$= (12 \times monthly\ factor) \times (3 \times daily\ factor) \times (24 \times hourly\ factor)$$

Where the factors were computed as follows:

Monthly factors computed as:

$$monthly\ factor = \frac{monthly\ emissions}{annual\ emissions}$$

Daily factors computed as:

$$weekday\ factor = \frac{average\ of\ weekday\ emissions\ for\ all\ 12\ months}{sum\ of\ 7\text{-}day\ emissions\ for\ all\ 12\ months}$$

$$Saturday\ factor = \frac{average\ of\ Saturday\ emissions\ for\ all\ 12\ months}{sum\ of\ 7\text{-}day\ emissions\ for\ all\ 12\ months}$$

$$Sunday\ factor = \frac{average\ of\ Sunday\ emissions\ for\ all\ 12\ months}{sum\ of\ 7\text{-}day\ emissions\ for\ all\ 12\ months}$$

Hourly factors computed as:

$$weekday\ hour\text{-}1\ factor = \frac{sum\ of\ hour\text{-}1\ for\ all\ weekdays}{sum\ of\ 24\text{-}hour\ emissions\ for\ all\ weekdays}$$

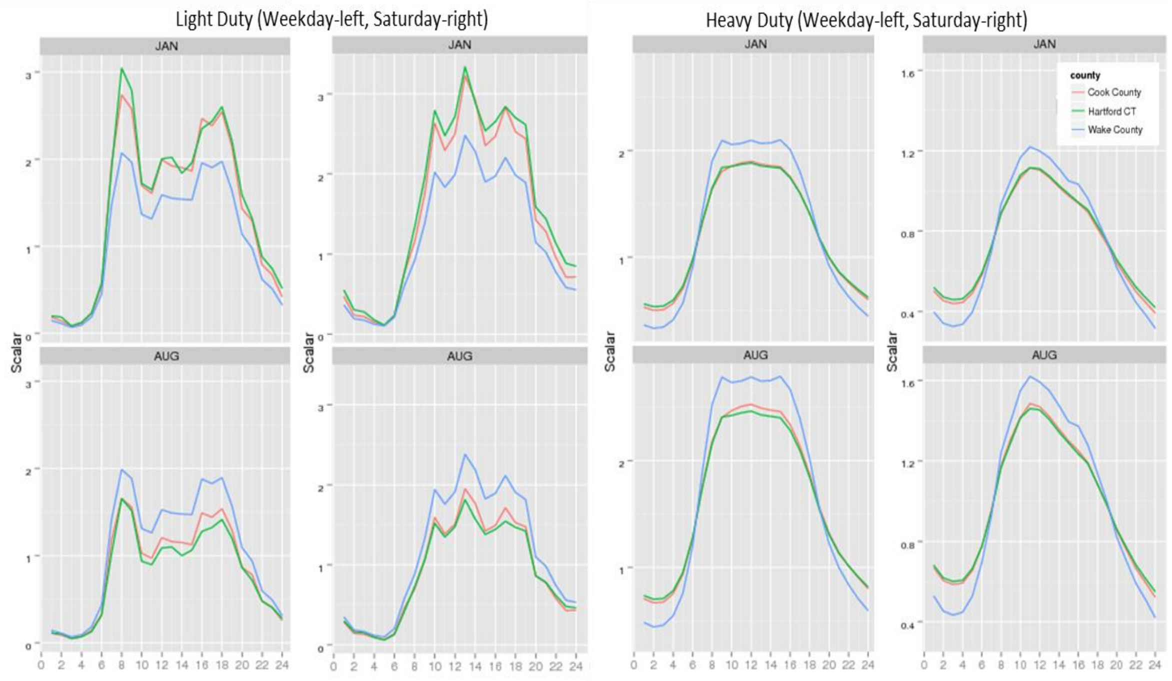
$$Saturday\ hour\text{-}1\ factor = \frac{sum\ of\ hour\text{-}1\ for\ all\ Saturdays}{sum\ of\ 24\text{-}hour\ emissions\ for\ all\ Saturdays}$$

$$Sunday\ hour\text{-}1\ factor = \frac{sum\ of\ hour\text{-}1\ for\ all\ Sundays}{sum\ of\ 24\text{-}hour\ emissions\ for\ all\ Sundays}$$

(etc.)

Exhibit 30 contains plots of temporal scalars for three counties, for January and August and a weekday versus a Saturday.

Exhibit 30. Example of Temporal Scalars by Hour-of-day for Onroad HEM Run Groups



2.4 Source Groups

Source attribution is a key feature of NATA because it allows users to prioritize sources and pollutants of interest. The source groupings for each NATA are a little different, in part due to methods that have increased the number of source groups allowable, and also due to the results of previous NATAs. For the 2011 NATA, we considered the NEI sectors, source groups developed for previous NATAs, and sources whose impacts were generated solely through CMAQ. Other than secondary formation and background, source groups originated by aggregating sources in the emission inventory based on attributes such as SCC and facility source type. The secondary-formation source group was based on the CMAQ estimates of the concentration of HAPs (formaldehyde, acetaldehyde, and acrolein) formed from the atmosphere due to photochemical reactions of anthropogenic and biogenic VOCs and NO_x. Background is the sum of two components:

- Impact of the remote-concentration estimates added to the HEM-3 runs for non-CMAQ HAPs (such as carbon tetrachloride) and non-CMAQ areas (i.e., outside the CONUS CMAQ domain)—see Section 3.2.1.
- Impact of the CMAQ concentration estimates in areas in which all HEM-3 receptors were zero but CMAQ were nonzero.

The source groups are provided in Exhibit 31.

Exhibit 31. Source Groups for NATA

Broad Group	NEI Data Category	NATA Source Group	Description	HEM Run Group	Based On
Point	Point	Point (PT)	Point sources excluding airports and railyards	Point—no airports	Point NEI data category
Onroad	Onroad	OR-LD_Gas	Onroad light-duty gasoline	Light Duty	SCC
		OR-HD_Gas	Onroad heavy-duty gasoline	Heavy Duty	
		OR-LD_Diesel	Onroad light-duty diesel	Light Duty	
		OR-HD_Diesel	Onroad heavy-duty diesel	Heavy Duty	
		OR-Refueling	Onroad refueling (including heavy- and light-duty gasoline and diesel refueling)	Light Duty	
Nonroad	Point	NR-Airport (PT)		Point—airports	Facility source type = 100
		NR-Railyard (PT)		Point—no airports	Facility source type = 151
	Nonpoint	NR-CMV Underway	CMVs Underway	CMVs	SCC
		NR-CMV Ports	CMVs at ports		
		NR-Locomotives		Nonpoint—10-m release height	
	Nonroad	NR-Pleasurecraft	Pleasurecraft (nonroad equipment)	Nonroad	SCC
		NR-Construction	Construction equipment		

Broad Group	NEI Data Category	NATA Source Group	Description	HEM Run Group	Based On
		NR-Gas/Other	Nonroad gasoline equipment other than pleasurecraft and construction		
		NR-Diesel/Other	Nonroad diesel equipment other than pleasurecraft and construction		
Nonpoint	Nonpoint	NP-Bulk_gas_term	Nonpoint bulk gasoline terminals	Nonpoint—10-m release height	SCC
		NP-Chemical_Mfg	Nonpoint chemical manufacturing		
		NP-Mining	Nonpoint mining		
		NP-Industrial_NEC	Nonpoint industrial processes not elsewhere classified		
		NP-Nonferrous_metals	Nonpoint nonferrous metals		
		NP-Oil/Gas	Nonpoint oil and gas production		
		NP-Refineries	Nonpoint refineries		
		NP-Storage_Transfer	Nonpoint storage and transfer		
		NP-ICI_fuel_comb	Nonpoint industrial, commercial, and institutional emissions		
		NP-Landfills	Nonpoint landfills [a]		
		NP-SfcCoating_IndSolvent	Nonpoint surface coating and industrial solvent use		
		NP-WasteDisposal_Other	Nonpoint other waste disposal		
		NP-Gas_stations	Nonpoint gas stations [b]	Nonpoint—low release height	
		NP-Comm_cooking	Nonpoint commercial cooking		
		NP-Misc_non-ind	Nonpoint miscellaneous nonindustrial sources		
		NP-Non-RWC_ResFuelComb	Nonpoint residential fuel combustion excluding wood		
		RWC	Nonpoint RWC		
		NP-Consumer_comm_solvent	Nonpoint consumer and commercial solvents		
		NP-Solvent_degreasing	Nonpoint solvent degreasing		
		NP-Dry_cleaning	Nonpoint dry cleaning		
		NP-Non-ind_sfc_coating	Nonpoint non-industrial surface coating		
Biogenics	Nonpoint	Biogenics		Not run in HEM-3	Zero out of CMAQ run

Broad Group	NEI Data Category	NATA Source Group	Description	HEM Run Group	Based On
Fires	Event	Fires (ag, prescribed and wild)	Fires - includes agricultural burning, prescribed burning and wildfires		Zero out of CMAQ run
Secondary	N/A	Secondary	Secondary transformation from anthropogenic and natural (i.e. all) sources		Zero out of CMAQ run
Background	N/A	Background			[c]

[a] Most landfills were in the point inventory in the 2011 NEI.

[b] Excluding Stage 2, which was covered in refueling.

[c] Non-CMAQ pollutants/areas: remote concentrations (e.g., carbon tetrachloride)—See Appendix E. For CMAQ pollutants/areas, background was only computed as the CMAQ concentration for grid cells with no HEM-3 contribution.

2.5 Uncertainties in Emissions/Emissions Processing

Uncertainties in emissions result from uncertainties in locations and release characteristics of sources, emission estimation techniques such as emission factors and models, inputs to emission models such as NMIM or MOVES, speciation factors, and activity data. For nonpoint, onroad, and nonroad categories, additional uncertainty lies with the spatial resolution of the activity data. For example, MOVES and NMIM estimated inventories using combined data for specific vehicles, activities, and fuels from states and/or EPA and other government agencies along with vehicle- and engine-emissions data. Activity data for vehicles and nonroad equipment were typically available at the levels of national, state, or metropolitan statistical area, and thus must be allocated to counties using surrogates such as population and land use. This allocation introduces significant uncertainty to county-level estimates of emissions.

RWC data were also estimated at the county level for all appliance types. For fireplaces and wood stoves, these allocations and burn rates were primarily based on American Housing Survey (AHS) data where available. AHS data was available for 47 select metropolitan areas every 6 years and assumptions were made based on survey-response rate, fraction of households that burn wood versus gas for primary and secondary heating, urban versus suburban and rural representation of wood-burning device-type use and burn rates. All of these characteristics have uncertainty and EPA used more local state and regional-level survey data where available. For other devices such as outdoor wood boilers (OWB) and indoor furnaces, national sales data were extrapolated from growth estimates to estimate total appliances. Spatial allocation utilized inverse population density with thresholds to prevent assigning OWBs to highly populated urban counties. Similar burn-rate zero-outs were applied for high-population-density counties for primary heating from most types of RWC devices. Each assumption in burn rate and appliance profile contained a varying level of uncertainty because of the sparse availability of survey data at the county level for each RWC device type.

Canada and Mexico did not have HAP emissions in the inventories except where VOC was speciated into benzene, formaldehyde, acetaldehyde, and methanol, which were model species for CB05. For Mexico, onroad VOC emissions tended to be overestimated, which caused very high benzene estimates at the border with Texas. This resulted in unreasonably high tract concentrations in Maverick County. To address this, we used the non-hybrid approach for the onroad HEM run groups for this county.

We did not include mercury emissions from Canada or Mexico (both had been included in the 2005 platform modeling).

2.6 Summary

- The 2011 NEI v2 was the primary source of emissions data for NATA, although some data were modified based on: comments received during the NATA review period, the requirements of the air quality models, and the available data on pollutant toxicity.
- Emissions were grouped into various categories of sources to track source contributions and to conform to the requirements of the air quality modeling methods.
- Emissions were processed separately for CMAQ and HEM-3 which have different temporal, spatial and speciation requirements for the input emissions.
- Emissions were allocated spatially based on various demographic, activity, and land-use surrogates.
- Emissions were allocated temporally primarily using established temporal factors.

3 AIR QUALITY MODELING & CHARACTERIZATION

The NATA emission estimates described in Section 2 were used as inputs to EPA air quality models to estimate ambient concentrations of emitted air toxics. An air quality model is a set of mathematical equations that uses emissions, meteorological data, and other information to simulate the behavior and movement of air toxics in the atmosphere. The air quality models used for NATA estimated outdoor concentrations of air toxics at specified locations. The NATA approach included the development and application of a new **hybrid approach** blending a **chemical transport model (CMAQ; EPA 2015g)** with a **dispersion model (HEM-3 running AERMOD**; see also the HEM-3 User's Guides, EPA 2014e**)** to estimate ambient concentrations of **40 of the more prevalent and higher risk HAPs** as described in Section 3.1.1. The air toxics modeled in the hybrid approach captured approximately 99 percent of the total risk nationally. The remaining "non-hybrid" air toxics were treated similarly to the NATA 2005 approach, which merged HEM-3 model estimates with observations, as described in Section 3.1.2.

For simplicity and consistency throughout this TSD, all aspects or details of the HEM-3 model are referred to overall as "HEM-3," although most often the AERMOD component of HEM-3 is pertinent to the discussion. EPA designed and maintains AERMOD separate and apart from HEM-3; HEM-3 merely incorporates AERMOD.

3.1 Hybrid Model Description

The subsections below contain discussions on the hybrid air modeling approach developed for the 2011 NATA.

3.1.1 Overview

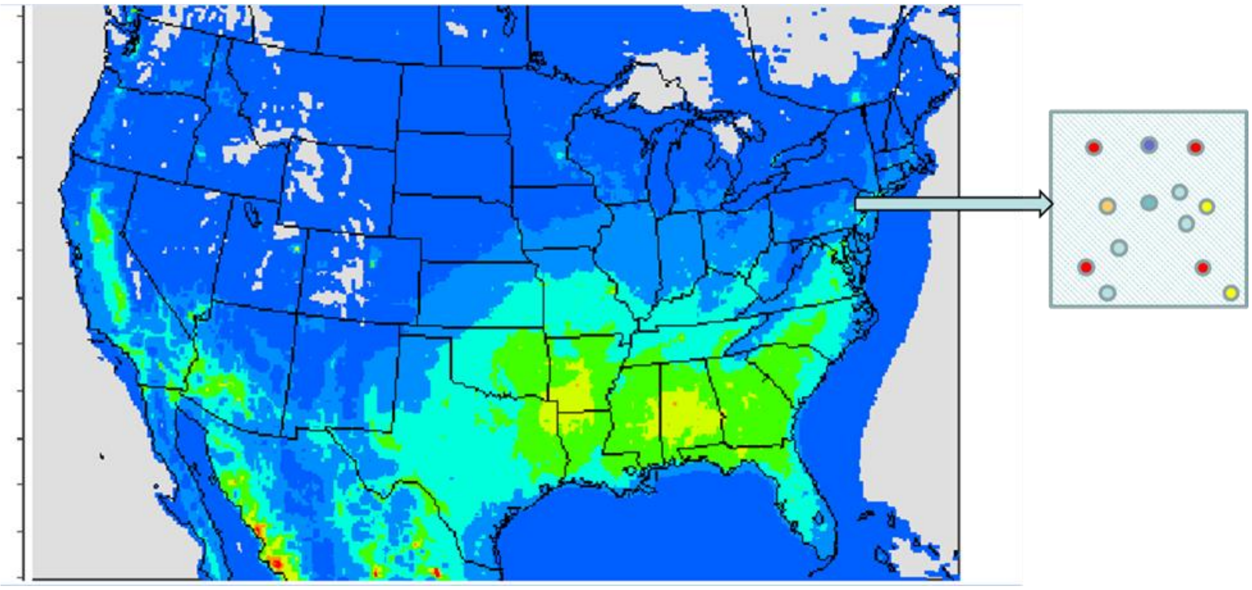
The air quality modeling structure for 40 of the most prevalent and highest risk air toxics (see Exhibit 32) utilized a hybrid method combining the fine spatial scale and source attributions of AERMOD in HEM-3 (Cimorelli et al. 2005; EPA 2015f) with the full treatment of chemistry and transport afforded by CMAQ version 5.02 with CB05 (Byun and Schere 2006; Brown et al. 2011). The CMAQ model is a comprehensive, three-dimensional grid-based Eulerian air quality model designed to simulate the formation and fate of gaseous and particulate species, including ozone, oxidant precursors, primary and secondary PM concentrations, and sulfur and nitrogen deposition over urban and regional spatial scales. In this application, HEM-3 treated all species as chemically non-reactive. The emissions and meteorological data sets used in CMAQ were processed further to generate HEM-3 inputs consistent with CMAQ. HEM-3 receptor locations were based on the centroids of populated census blocks, monitoring-site positions, and 5 evenly distributed points within each 12-km horizontal CMAQ grid cell in the CONUS (see Exhibit 33), resulting in at least 5, and sometimes more than 10,000, receptors per cell and 6.5 million receptors nationwide.

Exhibit 32. Air Toxics Utilizing the Hybrid Modeling in NATA

1,1,2,2-Tetrachloroethane	Benzene	Formaldehyde	Nickel Compounds
1,3-Butadiene	Beryllium	Hexamethylene	Nickel Compounds
1,3-Dichloropropene	Cadmium	Hydrochloric acid	Propylene dichloride
1,4-Dichlorobenzene(p)	Chlorine	Lead Compounds	Quinoline
2,4-Toluene diisocyanate	Chloroform	Maleic anhydride	Tetrachloroethylene

- Acetaldehyde
- Acetonitrile
- Acrolein
- Acrylonitrile
- Arsenic
- Chromium (VI)
- Diesel PM
- Ethylene dibromide
- Ethylene dichloride
- Ethylene oxide
- Manganese Compounds
- Mercury Compounds
- Methanol
- Methylene chloride
- Naphthalene
- Toluene
- Trichloroethylene
- Triethylamine
- Vinyl chloride
- Xylenes

Exhibit 33. CMAQ Domain with Expanded Cell Showing Hybrid Receptors



The equation below was used to calculate the 2011 annual-average estimates of air concentrations at receptor locations, which were constrained to CMAQ-grid-average values, with HEM-3 providing sub-grid scale spatial texture.

$$C = HEM\text{-}3_{REC} \times \left(\frac{CMAQ_{PNFB}}{HEM\text{-}3_{GRIDAVG}}\right) + CMAQ_{SEC} + CMAQ_{PFIRES} + CMAQ_{PBIOGENICS}$$

Where:

C = concentration at a receptor,

$CMAQ_{PNFB}$ = concentration in CMAQ grid cell, contributed by primary emissions, excluding fires and biogenics,

$HEM\text{-}3_{REC}$ = concentration at HEM-3 receptor,

$HEM\text{-}3_{GRIDAVG}$ = average of all HEM-3 results within a CMAQ grid, calculated through surface interpolation of all HEM-3 receptor locations to eliminate concentration discontinuities,

$CMAQ_{SEC}$ = contribution from atmospheric reactions in CMAQ grid cell,

$CMAQ_{PFIRES}$ = contribution from primary emissions of fires in CMAQ grid cell, and

$CMAQ_{PBIOGENICS}$ = contribution from primary emissions of biogenics in CMAQ grid cell.

This hybrid approach, which builds on earlier area-specific applications to Philadelphia, PA (Isakov et al. 2007) and Detroit, MI (Wesson et al. 2010), reflects an evolution of national-scale modeling of HAPs that attempts to optimize characterization of non-reactive and reactive species across multiple spatial scales. However, the blending of two different modeling platforms challenges adherence to basic mass-conservation principles. For example, the 2005 NATA was susceptible to duplicate counting as the secondarily formed species generated by CMAQ were added directly to HEM-3 estimates driven by primary emissions, in addition to an added "background" concentration based on ambient observations. Consequently, diagnosing model behavior based on paired model-to-measurement values was compromised by the dual use of observations and largely inconsistent model inputs driving HEM-3 and CMAQ. CMAQ tracks primary and secondary contributions by source type, enabling the HEM-3 estimate at each receptor location to be normalized to the CMAQ primary contribution. By **anchoring concentration averages to CMAQ**, mass conservation was largely retained. The constraint to CMAQ average grid values imposed by the above equation minimized possible redundancies and was appropriate when combining results from vastly different model architectures.

3.1.2 Treatment of Species

In this application, the hybrid model was applied to 40 of the highest risk air toxics (shown in Exhibit 32) among 180 air toxics included in the 2011 NATA. Although this application focuses on air toxics, it reflects the second major application of the CMAQ multipollutant version incorporating CAPs and HAPs following a national assessment of increased ethanol use associated with renewable fuels (Cook et al. 2011). The atmospheric chemistry treatments in chemical transport models such as CMAQ are based on gas-phase reaction processes optimized to characterize ozone, linked with a variety of heterogeneous and thermodynamic processes to accommodate PM formation. Consequently, the inclusion of explicit chemical species in current chemical mechanisms is predicated by its relative importance in ozone chemistry. Formaldehyde and acetaldehyde are high risk HAPs that generate significant amounts of peroxy radicals leading to enhanced ozone production and secondary PM formation and exemplify multipollutant linkages driven by atmospheric processes. Chemical species that are not incorporated as explicit species in chemical mechanisms are added as non-reactive tracers (e.g., several halogenates) or included in simple reaction schemes, such as 1,3-butadiene decay and subsequent acrolein generation, decoupled from the chemical mechanism. The emissions mass of several less reactive VOCs such as the prevalent benzene, toluene, and xylene species are tracked as non-reactive tracers and participate in atmospheric reactions as lumped carbon bond species with the assumption that atmospheric chemistry minimally influences air concentrations. HEM-3, which treats all pollutants as nonreactive, was applied to the remaining air toxics not incorporated within CMAQ. Due to relatively spatially invariant concentration distributions and uncertainty in emission estimates, carbon tetrachloride risk was based on observations modulated spatially by HEM-3 estimates.

3.1.3 Meteorological Processing

The gridded meteorological data for 2011 at the 12-km CONUS-scale domain (see domain in Exhibit 33) was derived from version 3.4 of **WRF** (Skamarock et al. 2008). The WRF meteorological outputs were processed using the Meteorology-Chemistry Interface Processor package (version 4.1.3) to derive hourly specific inputs to CMAQ: horizontal wind components (i.e., speed and direction), temperature, moisture, vertical-diffusion rates, and rainfall rates for each grid cell in each vertical layer (Otte and Pleim 2010). A performance evaluation of the meteorological model can be found in EPA (2014a). CMAQ resolved the vertical atmosphere with 25 layers, preserving greater resolution in the planetary boundary layer. The meteorological inputs driving CMAQ were processed for HEM-3 through the Mesoscale Model Interface Program (MMIF; EPA 2014d) to provide representative meteorological inputs for every fourth CMAQ grid cell and for over 700 National Weather Service station locations. Isakov et al. (2007) demonstrated successful use of meteorological variables derived from prognostic modeling to drive dispersion models, motivated primarily to address spatial gaps in meteorological monitoring. Additionally, MMIF outputs have been shown to compare favorably well against observed meteorological data when used in AERMOD (EPA 2015m). For the 2011 NATA, we took advantage of **prognostic meteorological modeling** to harmonize boundary-layer inputs across both modeling platforms. The MMIF outputs were based on the first CMAQ application using the 2011 NEI v1. The final CMAQ application, based on the 2011 NEI v2, included a new WRF simulation that was not processed further to update MMIF outputs. The differences in these WRF simulations were insignificant with respect to affecting the MMIF results.

3.1.4 Emissions Processing Overview

See Section 2 for additional detail. The 2011 NEI provided the root emissions data for CMAQ and HEM-3 (EPA 2014b). Emissions to HEM-3 were grouped into four broad categories (emissions-input resolution of over 150 SCCs was retained) with similar spatial and temporal delineation: **major point sources**, **nonpoint sources** (excluding transportation, fires, and biogenics), **onroad mobile**, and **nonroad mobile** (including locomotive, aviation, and CMVs). **Fires** (combined wild and prescribed) and **biogenic** emissions were handled only through CMAQ. NEI data were provided as specific point and aggregated county-level annual estimates and were processed to hourly values distributed over 12-km horizontal grids through SMOKE (Houyoux et al. 2000). **Hourly temporal allocations** were developed for HEM-3, consistent with CMAQ, which was a departure from previous NATAs that were based on annual-average inputs. Nonpoint, onroad, and nonroad emissions for HEM-3 typically were allocated spatially to population census tracts using a variety of surrogates (e.g., land use classifications, population).

3.1.5 Initial and Boundary Conditions

The CMAQ lateral-boundary and initial-species concentrations for benzene, formaldehyde, and acetaldehyde were generated by a year-2011GEOS-Chem simulation (Yantosca et al. 2015). The 2011 GEOS-Chem simulation used a grid resolution of 2.0 degrees x 2.5 degrees (latitude x longitude). The predictions were processed using the GEOS-2-CMAQ tool and used to provide one-way dynamic boundary conditions at one-hour intervals (Akhtar et al., 2012). A GEOS-Chem evaluation was conducted for the purpose of validating the 2011 GEOS-Chem simulation for predicting selected measurements relevant to their use as boundary conditions for CMAQ. This evaluation included using satellite retrievals paired with GEOS-Chem grid cells (Henderson, et al., 2014). Due to the scarcity of observations suitable for establishing boundary conditions and the extended calendar year simulation, zero-value initial and boundary conditions were used for the remaining air toxics.

3.1.6 Source Attribution

Since the HEM-3 concentration estimates were based on separate emissions-group runs, estimates of the source contributions associated with primary emissions were generated by the following ratio technique normalized to CMAQ concentrations for sources within the four broad categories of sources:

$$C_{REC,J} = HEM\text{-}3_{REC,J} \times \frac{CMAQ_{PNFB}}{HEM\text{-}3_{GRIDAVG}}$$

Where:

$C_{REC,J}$ = contribution to concentration at a receptor REC from category J, excluding secondary formation.

This ratio approach provided an estimate of primary-emission contributions only. Primary-emission contributions from biogenics and fires were processed only through CMAQ, and all contributions from secondary formation processes were aggregated into $CMAQ_{SECNFB}$.

3.2 Treatment of Non-hybrid Air Toxics and Areas Outside the CONUS

The 138 HAPs not simulated by CMAQ were estimated by adding HEM-3-modeled concentrations to observed ambient concentrations assumed to reflect background conditions, utilizing the same emissions and meteorological processing used for HEM-3 in the hybrid modeling. The non-CONUS locations (Alaska, Hawaii, Puerto Rico, and U.S. Virgin Islands) were modeled with the same non-hybrid approach, but with all NATA air toxics included.

3.2.1 Background Concentrations

Appendix E contains further details on background concentrations. For the 2011 NATA, we added background concentrations to the HEM-3-modeled concentrations (i.e., those pollutants that were not modeled in CMAQ). Background concentrations were added to account for (1) natural sources, (2) emissions of persistent air toxics that occurred in previous years, and (3) long-range transport from distant sources. The largest impact of the background concentrations on the NATA risk was **carbon tetrachloride**, a ubiquitous, globally persistent HAP for which the background concentration of 0.55 $\mu g/m^3$ was 3.2 times higher than its URE of 0.17 $\mu g/m^3$, thus providing a nationally consistent background cancer risk of 3-in-1 million. Although no large sources of this pollutant exist because the Montreal Protocol established a timetable for its phase-out, observations in national and remote networks show fairly uniform concentrations, possibly partly due to its long residence time (30–50 years) or potentially resulting from unknown global sources.

Various approaches have been used in previous NATAs, including use of ambient concentration data, remote concentration data, and emissions. The approach for the 2011 NATA was simplified from previous approaches because the available data did not support the generation of regionally varying background concentrations. For all air toxics, we estimated a spatially uniform, remote Northern Hemisphere air toxics concentration for the background. The remote concentrations were estimated from measurements made at remote monitoring sites, from lower-level concentrations from the national monitoring network, from the scientific literature, or from application of national-emissions and residence-time ratios to a measured remote concentration.

The generalized methods hierarchy was as follows:

1. **Remote network**. Measurements made at networks/sites in remote Northern Hemisphere locations with citations in peer-reviewed literature. Examples include National Oceanic and Atmospheric Administration Global Monitoring Division sites, the Trinidad Head Advanced Global Atmospheric Gases Experiment site, and remote Interagency Monitoring of Protected Visual Environments (IMPROVE) sites.
2. **Ambient national network**. Estimates of background concentrations made at routine monitoring network sites in the United States where concentrations were measurable and reliable down to the 10th percentile. Air toxics measurements were used from Phase XIII of the air toxics archive.
3. **Literature**. Measurements of species at remote or regional sites during a single-measurement study for a specific pollutant. These measurements are different from the ongoing measurements used in Step 1 because of their "one-off" nature.

4. **Emissions**. Emissions-based estimates of remote background concentrations. An estimate for pollutant *i* was derived from the national total emissions from the 2011 NEI (E), atmospheric residence times (t), and a comparison to tetrachloroethylene (PCE). The below equation shows the relationship used to derive these remote estimates.

$$C_i = \frac{E_i \times t_i \times C_{PCE}}{E_{PCE} \times t_{PCE}}$$

For NATA, we do not use background concentrations for formaldehyde, acetaldehyde, or acrolein because Alaska, Hawaii, Puerto Rico, and the U.S. Virgin Islands used national-average secondary concentrations from the CMAQ run, and adding background for these HAPs would be double counting. Also, we did not use the estimates for PAHs due to uncertainties in the residence times.

Exhibit 34 contains the background concentrations for the non-CMAQ air toxics, used for both the CONUS and non-CONUS areas, and the method used to generate them. Residence-time data were used only for the emissions-based approach. In addition to the background concentrations in Alaska, Hawaii, Puerto Rico, and the U.S. Virgin Islands also used the background concentrations for the pollutants listed in Exhibit 35. Diesel PM does not have a background concentration.

Exhibit 34. Background Concentrations Added to the HEM-3 Concentrations for Non-CMAQ Air Toxics, All Areas

Pollutant	Residence Time (days)	2011 Remote Concentration Estimate (µg/m^3)	Method	Year	Locations
Methyl chloride (Chloromethane)	365	1.09	Remote Network	2011	Trinidad Head, CA
Carbon tetrachloride	10950	0.547	Remote Network	2011	KUM, MLO, NWR, BRW, ALT
Acetonitrile	365	0.170	Literature	2004-2008	New Hampshire
n-Hexane	2.6	0.114	Emissions	2011	
2,2,4-Trimethylpentane	4	0.093	Emissions	2011	
1,1,1 Trichloroethane (Methyl chloroform)	1825	0.06	Remote Network	2011	KUM, MLO, NWR, BRW, ALT
Methyl bromide	365	0.029	Remote Network	2011	KUM, MLO, NWR, BRW, ALT
Ethylbenzene	1.7	0.016	Literature	2005	Jungfraujoch, Switzerland
Bromoform	540	0.01	Literature	1994-2004	Ocean cruises
Bis(2-ethylhexyl) phthalate	200	5.7E-03	Emissions	2011	
Carbon disulfide	7	5.5E-03	Emissions	2011	
Methyl isobutyl ketone	1	4.6E-03	Emissions	2011	
Styrene	0.25	8.7E-04	Emissions	2011	
1,1,2-Trichloroethane	49	3.9E-04	Emissions	2011	

Pollutant	Residence Time (days)	2011 Remote Concentration Estimate (µg/m^3)	Method	Year	Locations
Cumene (Isopropylbenzene)	2.2	3.7E-04	Emissions	2011	
Selenium	10	2.0E-04	Air Toxics Archive	2010-2012	All national sites
Benzyl chloride	3	1.4E-04	Emissions	2011	
Propionaldehyde (Propanal)	0.1	9.8E-05	Emissions	2011	
Antimony	10	6.6E-05	Emissions	2011	
Cobalt	10	4.1E-05	Air Toxics Archive	2010-2012	All national sites
Vinyl acetate	0.25	3.5E-05	Emissions	2011	
1,2-Dibromo-3-chloropropane	36	1.3E-06	Emissions	2011	
Benzidine	1	2.5E-08	Emissions	2011	

Exhibit 35. Background Concentrations Added to the HEM-3 Concentrations for Non-CONUS Areas Only

Name	Residence Time (days)	Best 2011 RCE Estimate (µg/m^3)	Method	Year	Locations
Dichloromethane (Methylene Chloride)	30	0.146	Remote Network	2011	KUM, MLO, NWR, BRW, ALT
Benzene	3	0.116	Remote Network	2011	KUM, MLO, NWR, BRW, ALT
Chloroform	80	0.058	Remote Network	2011	Trinidad Head, CA
Toluene	0.5	0.041	Literature	2005	Jungfraujoch, Switzerland
Xylenes	0.2	0.023	Literature	2002	Trinidad Head, CA
Tetrachloroethylene	6.5	0.013	Remote Network	2011	KUM, MLO, NWR, BRW, ALT
1,4-Dichlorobenzene	31	9.3E-3	Emissions	2011	
Trichloroethylene	6	4.1E-03	Remote Network	2011	Trinidad Head, CA
Ethylene dichloride	42	2.0E-03	Emissions	2011	
1,3-Butadiene	0.08	2.0E-03	Literature	2011	Jungfraujoch, Switzerland
1,1,2,2-Tetrachloroethane	91.3	8.4E-04	Emissions	2011	
Lead	10	6.6E-04	Remote Network	2010-2012	DENA, KALM, PORE, REDW, TRCR, TUXE,HACR
1,3-Dichloropropene	1.25	6.2E-04	Emissions	2011	

Name	Residence Time (days)	Best 2011 RCE Estimate (µg/m^3)	Method	Year	Locations
Naphthalene	0.25	4.9E-04	Emissions	2011	
Ethylene oxide	7	3.4E-04	Emissions	2011	
Manganese	10	3.2E-04	Remote Network	2010-2012	DENA, KALM, PORE, REDW, TRCR, TUXE,HACR
Acrylonitrile	5.6	3.2E-04	Emissions	2011	
1,2-Dichloropropane (propylene dichloride)	30	2.9E-04	Emissions	2011	
Ethylene dibromide	50	1.9E-04	Emissions	2011	
Arsenic	10	1.4E-04	Air Toxics Archive	2010-2012	All national sites
Vinyl chloride	2	1.2E-04	Emissions	2011	
Nickel	10	1.0E-04	Remote Network	2010-2012	DENA, KALM, PORE, REDW, TRCR, TUXE,HACR
Mercury (gaseous)	365	1.0E-04	Literature	2005	U.S. MDN sites
Cadmium	10	3.8E-05	Air Toxics Archive	2010-2012	All national sites
Beryllium	10	8.5E-06	Air Toxics Archive	2010-2012	All national sites
Chromium VI	3	1.5E-06	Remote Network and National Measurements	2010-2012	All national sites
Hydrazine	0.25	5.7E-8	Emissions	2011	

For the 2011 NATA, two pollutants from the 2002 and 2005 NATA background lists were excluded: quinoline and total chromium (although chromium (VI) was included in 2011 NATA). Four pollutants (or grouped pollutants) that had been in the 1999 NATA background list were excluded: polychlorinated biphenyls, lindane, phosgene, and hexachlorobutadiene. In contrast, some pollutants and groups that had never previously been estimated were added—these include acetonitrile, hexane, 2,2,4-trimethylpentane, ethylbenzene, methyl isobutyl ketone, styrene, 1,3-dichloropropene, 1,1,2-trichloroethane, cumene(isopropylbenzene), selenium, propionaldehyde (propanal), antimony, cobalt, and vinyl acetate.

3.3 Model Evaluation

We describe below and in Appendix F our efforts to evaluate the performance of the NATA models. As noted at the beginning of Section 3, discussions of "HEM-3" in this document often are specifically related to the AERMOD dispersion model component of HEM-3, but we use "HEM-3" throughout for simplicity and consistency. In this section in particular, discussions of HEM-3 model values are specifically related to the air concentrations predicted by its AERMOD component.

3.3.1 Overview

Using the air toxics archive Phase IX for the year 2011, we conducted an operational model performance evaluation of the air toxics simulated for the 2011 NATA (more details found in Section 3.3.2 below). The model evaluation included both the air toxics modeled with the hybrid approach ("hybrid air toxics") and those modeled without the hybrid approach ("non-hybrid air toxics"). The hybrid evaluation looked at the air toxics for which there were valid ambient data (i.e., completeness criteria protocol) to compare against the CMAQ, HEM-3, and hybrid model predictions. Likewise, the air toxics non-hybrid evaluation used similar observational-completeness criteria constraints to compare against air toxics estimated by adding HEM-3 to observed ambient concentrations assumed to reflect background conditions.

Spatial-scale differences exist between CMAQ, HEM-3, and the hybrid model predictions. A CMAQ concentration represents a 12-km grid-cell volume-averaged value. The HEM-3 model concentration represents a specific point within the modeled domain. The hybrid model concentration combines the HEM-3 point-concentration gradients with the CMAQ 12-km grid-cell volume average. The ambient observed measurements were made at specific spatial locations (latitude/longitude). Several annual graphical presentations and statistics of model performance were calculated and prepared. Graphical presentations included box and whisker plots (which show the distribution and the bias of the predicted and observed data) and regional maps (which show the mean bias and error calculated at individual monitoring sites).

3.3.2 Observations

Observations were extracted from the air toxics archive, Phase IX for the year 2011. While most of the data in the archive are a snapshot of the Air Quality System (AQS) database (downloaded in July 2014), additional data (such as from special studies) were in the archive but not reported to AQS. In the air toxics archive, pollutant concentrations were converted to $\mu g/m^3$ in local conditions where temperature and pressure data were available (i.e., at the vast majority of sites). In addition, any negatives and data flagged as "non-detect" without a value were given a value of 0. Also, any data determined to have been substituted with half the method detection limit (i.e., MDL/2) was changed to 0.

For comparing annual averages of modeled and monitored data, data from the archive were aggregated to 2011 annual averages by site and parameter code. Data below MDLs were used as-is. Data were removed for which there were no MDLs. Naphthalene data from parameter code 45850 (canister method) were removed because that method may not be as reliable as the method used in the National Ambient Air Toxics Trends Program. Also, those sites were removed that reported naphthalene as code 17141 in which it was determined to use the canister test method.

Only site-parameter pairs in which measurements from at least three seasons were 75 percent complete (i.e., 75 percent of the scheduled days contained non-null values) were retained for developing annual averages. First, the sub-annual data were allocated to 24-hour averages. Seventy-five-percent completeness was required to create a daily average from sub-daily data, such that 75 percent of scheduled sub-annual data were available. For example, hourly data required 18 of 24 hours of data, three-hour data required six of eight three-hour periods of data, etc. For each quarter, the number of days to meet 75% completeness depended on the sampling frequency (note that more than one monitor at the site that measured on the same day was counted once). For example, one-in-six-day sampling required 12 days for the quarter.

For sites with multiple monitors (known as "POCS"), only the daily data with the same measurement duration (i.e., hourly, 3-hour, 24-hour) were averaged across the POCs. That is, daily data based on hourly measurements were not averaged with daily data based on 24-hour measurements. Where a site met the 75% completeness for multiple durations, the 24-hour duration data were chosen.

Annual averages were created by averaging all daily measurements with the same measurement duration for all sites that met the above completeness criteria. The only sites used were those for which 50 percent or more of the data were above the method detection limit (MDL). A spreadsheet file ("2011monitored_data_annualmeans_PhaseIXarchive.xlsx") of the ambient annual averages (in µg/m^3) is provided in the SupplementalData folder.

Uncertainties in the ambient data result from limited sites, data below MDL and measurement uncertainties.

3.3.3 Model Performance Statistics

The Atmospheric Model Evaluation Tool (AMET) was used to conduct the 2011 NATA air toxics evaluation (Appel et al. 2011). There are various statistical metrics available and used by the science community for model performance evaluation. For a robust evaluation, the principal evaluation statistics used to evaluate model performance are based on the following metrics: two bias metrics (mean bias and normalized mean bias); and, three error metrics (mean error and normalized mean error, root mean square error, and correlation coefficient).

Common variables are:

M = predicted concentration
O = observed concentration
X = predicted or observed concentration
σ = standard deviation

Mean Bias (MB), Mean Error (ME), and Root Mean Square Error (RMSE) (all in µg/m^3):

$$MB = \frac{1}{n}\sum_{1}^{n}(M - O)$$

$$ME = \frac{1}{n}\sum_{1}^{n}|M - O|$$

$$RMSE = \sqrt{\frac{\sum_{1}^{n}(M - O)^2}{n}}$$

MB quantifies the tendency of the model to over- or under-estimate values while ME and RMSE measure the magnitude of the difference between modeled and observe values regardless of whether the modeled values are higher or lower than observations.

Normalized Mean Bias (NMB) and Normalized Mean Error (NME) (both unitless):

$$NMB = \frac{\sum_{1}^{n}(M - O)}{\sum_{1}^{n} O}$$

NMB is used as a normalization to facilitate a range of concentration magnitudes. This statistic averages the difference (model - observed) over the sum of observed values. NMB is a useful model performance indicator because it avoids over-inflating the observed range of values, especially at low concentrations.

$$NME = \frac{\sum_1^n |M - O|}{\sum_1^n O}$$

NME is similar to NMB, where the performance statistic is used as a normalization of the ME. NME indicates the absolute value of the difference (model - observed) over the sum of observed values.

Correlation Coefficient (r; unitless):

$$r = \frac{1}{(n-1)} \sum_1^n \left[\left(\frac{O - \bar{O}}{\sigma_o} \right) \times \left(\frac{M - \bar{M}}{\sigma_m} \right) \right]$$

The value of r provides an indication of the strength of linear relationship and is signed positive or negative based on the slope of the linear regression.

3.3.4 Hybrid Evaluation

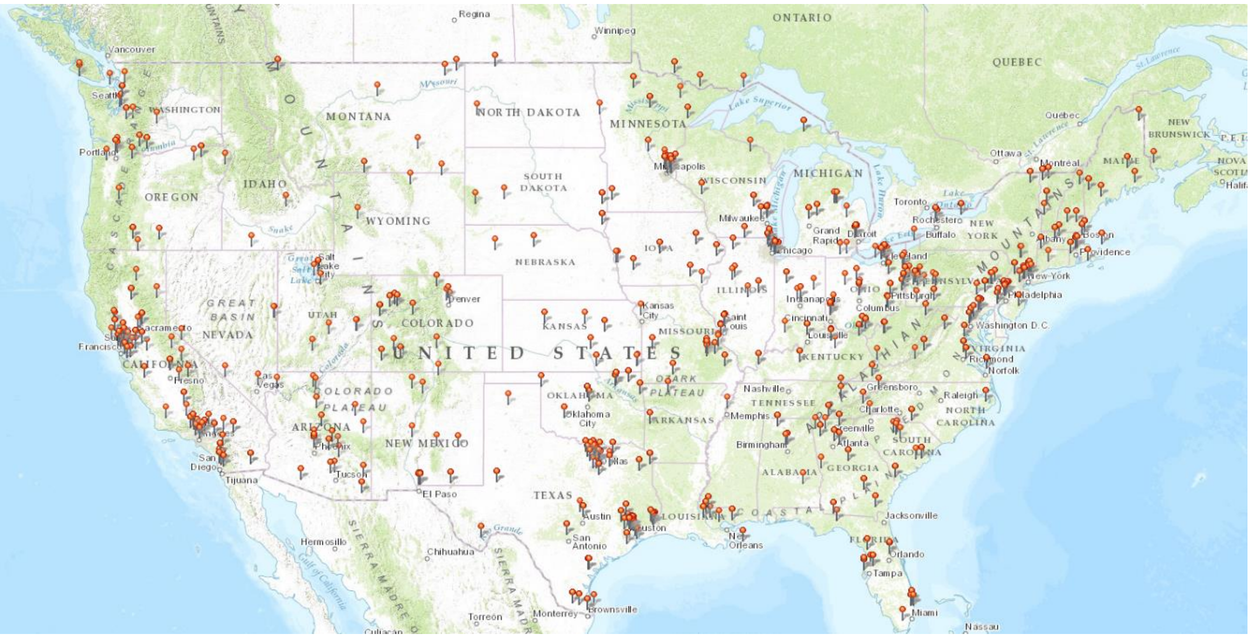
We conducted an annual operational model performance evaluation for hybrid air toxics, resulting in comparisons between CMAQ and HEM-3 predictions as well as an evaluation of the ability of the hybrid model to replicate the 2011 observed ambient concentrations. Inclusion of all three model results was intended to demonstrate the merged attributes of the hybrid model used for the 2011 NATA. Statistical assessments of modeled versus observed concentrations were paired in time and space and aggregated on an annual basis. Exhibit 36 contains a list of air toxics evaluated in the hybrid model performance evaluation and the number of paired sites (based on completeness criteria of observations, Section 3.3.2) used in the annual average. Exhibit 37 is a map of the 2011 monitoring locations for hybrid air toxics. Acrolein and ethylene dibromide were excluded in the model evaluation given the data uncertainty and sampling. Annual averages of xylene species (m-, o-, and p-) were summed together to calculate a "Xylenes, total" air toxic group based on the individual risk for each species being the same.

Exhibit 36. Hybrid Air Toxics Evaluated

Hybrid Air Toxic	Number of Paired Sites
1,3-Butadiene	83
1,3-Dichloropropene	5
1,4-Dichlorobenzene	22
Acetaldehyde	110
Acrylonitrile	18
Arsenic PM_{10}	34
Benzene	214
Cadmium PM_{10}	27
Chlorine	123
Chloroform	92
Ethylene dichloride	40

Hybrid Air Toxic	Number of Paired Sites
Formaldehyde	110
Lead PM_{10}	33
Manganese PM_{10}	40
Methylene chloride	123
Naphthalene	36
Nickel PM_{10}	29
Propylene dichloride	5
Tetrachloroethylene	72
Toluene	211
Trichloroethylene	13
Xylenes	163

Exhibit 37. 2011 Monitoring Locations for the Evaluation of Hybrid Air Toxics



In this section of evaluation of hybrid air toxics, we present annual-average model-to-monitor site comparisons for all three models for three key air toxics: acetaldehyde, formaldehyde, and benzene. Presentation of these comparisons includes summary statistics (Exhibit 38), boxplots containing model distribution and bias differences as compared to ambient observations (Exhibit 39, Exhibit 46, and Exhibit 53), and regional spatial maps with plots of the mean bias and error calculated at individual monitoring sites (Exhibit 40–Exhibit 45, Exhibit 47–Exhibit 52, Exhibit 54–Exhibit 59). The boxplots use boxed interquartile ranges of 25th-to-75th percentile, along with whiskers from the 5th to 95th percentiles, and they also contain summary statistics of r, RMSE, NMB, NME, MB, and ME. More details of the hybrid evaluation are in Appendix F, including statistical assessments of all the evaluated hybrid air toxics. Note that HEM-3 and AERMOD are equivalent terms in these exhibits.

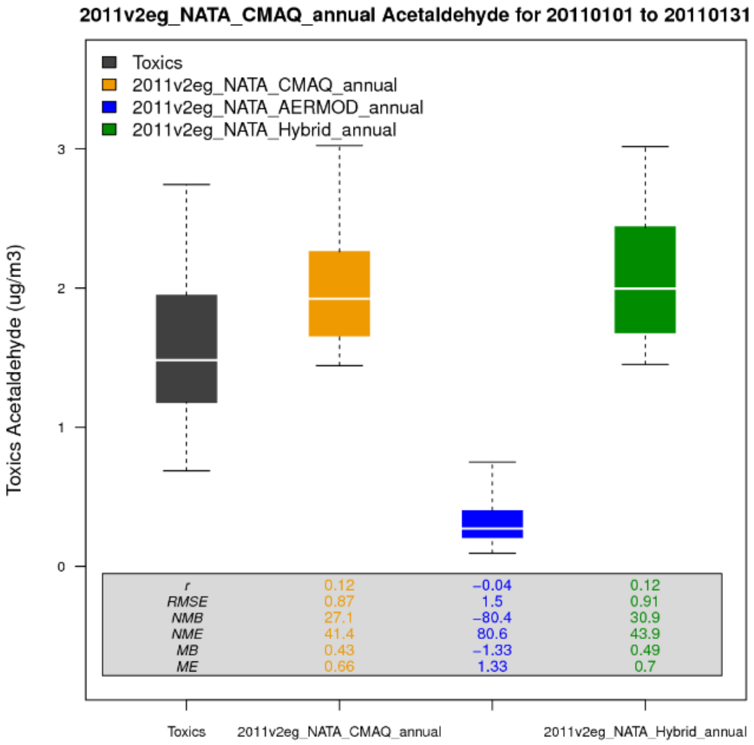
CMAQ and hybrid model predictions of annual formaldehyde, acetaldehyde, and benzene showed relatively small-to-moderate bias and error percentages when compared to observations. HEM-3 showed larger biases and errors, with underestimates for secondarily formed air toxics (e.g., -80.4 percent for acetaldehyde and -78.8 percent for formaldehyde), as expected given HEM-3's exclusion of atmospheric chemistry. Differences in bias and error statistics between the hybrid and CMAQ models were negligible for formaldehyde and acetaldehyde. Technical issues in the air toxics data consisted of (1) uncertainties in monitoring methods, (2) limited measurements in time/space to characterize ambient concentrations ("local in nature"), (3) commensurability issues between measurements and model predictions, (4) emissions- and science-uncertainty issues potentially affecting model performance, and (5) limited data for estimating intercontinental transport that effects the estimation of boundary conditions (i.e., boundary estimates for some species were much higher than predicted values inside the domain).

Exhibit 38. 2011 Annual Air Toxics Performance Statistics for the Hybrid, CMAQ, and HEM-3 Models

Hybrid Air Toxic	Model	MB (µg/m^3)	ME (µg/m^3)	NMB (%)	NME (%)
Acetaldehyde	Hybrid	0.5	0.7	30.9	43.9
	CMAQ	0.4	0.7	27.1	41.4

Hybrid Air Toxic	Model	MB (µg/m³)	ME (µg/m³)	NMB (%)	NME (%)
	HEM-3	-1.3	1.3	-80.4	80.6
Formaldehyde	Hybrid	-0.8	1.0	-30.8	37.3
	CMAQ	-0.9	1.0	-34.2	38.8
	HEM-3	-2.2	2.2	-78.8	79.3
Benzene	Hybrid	0.0	0.5	1.6	60.7
	CMAQ	-0.2	0.5	-22.7	57.8
	HEM-3	-0.3	0.5	-33.1	60.2

Exhibit 39. Acetaldehyde: 2011 Boxplots of Observed and Modeled Concentrations (top) and Modeled-Observed Bias Difference in Concentrations (bottom) for the Hybrid, CMAQ, and HEM-3 Models



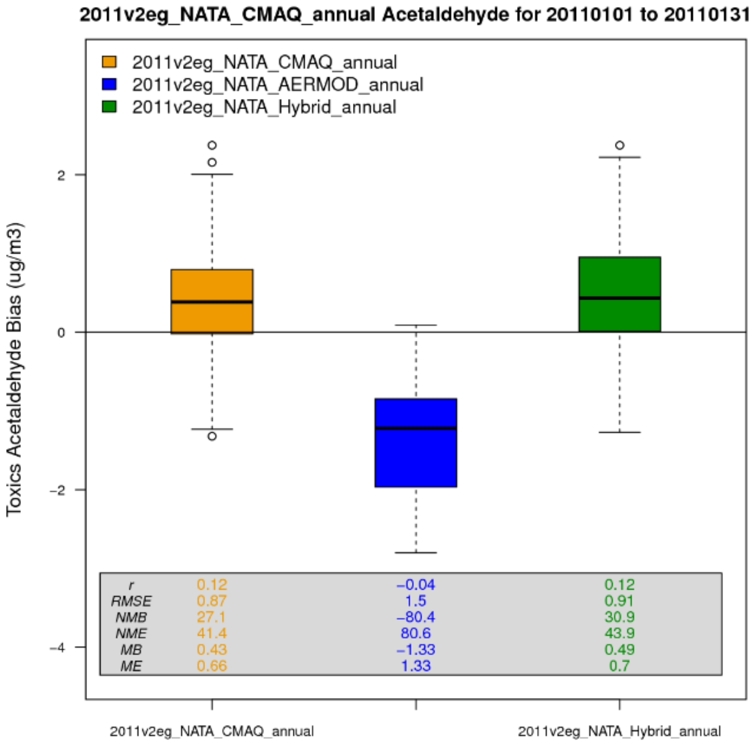


Exhibit 40. Acetaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the Hybrid Modeling Domain

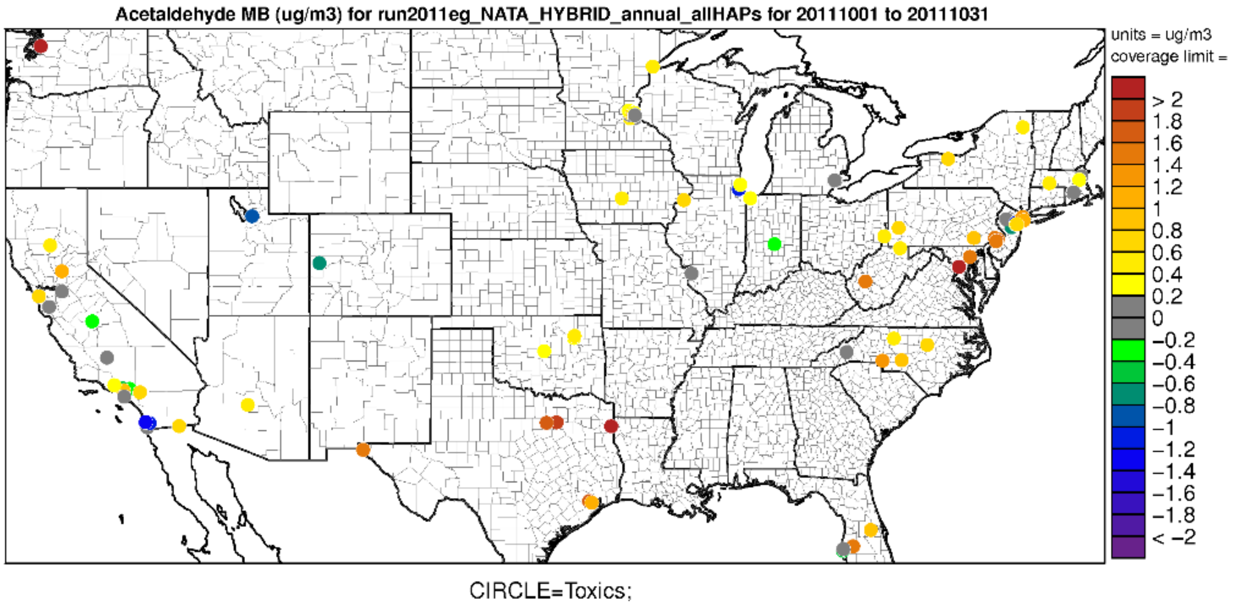


Exhibit 41. Acetaldehyde: 2011 Mean Error (%) at Monitoring Sites in the Hybrid Modeling Domain

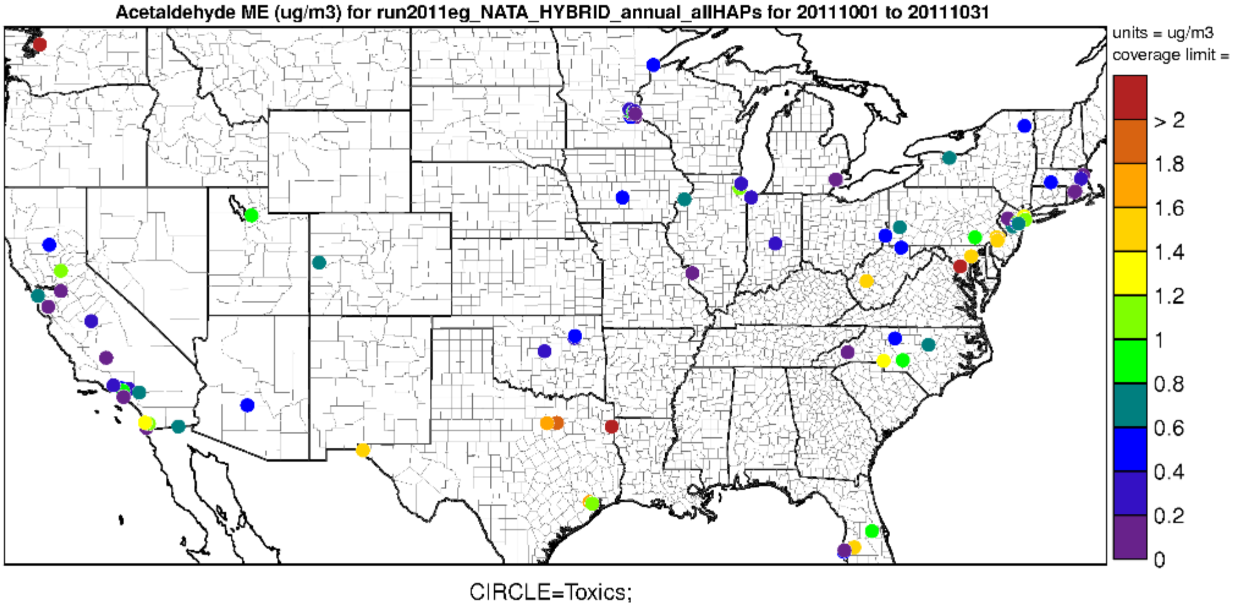


Exhibit 42. Acetaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the CMAQ Modeling Domain

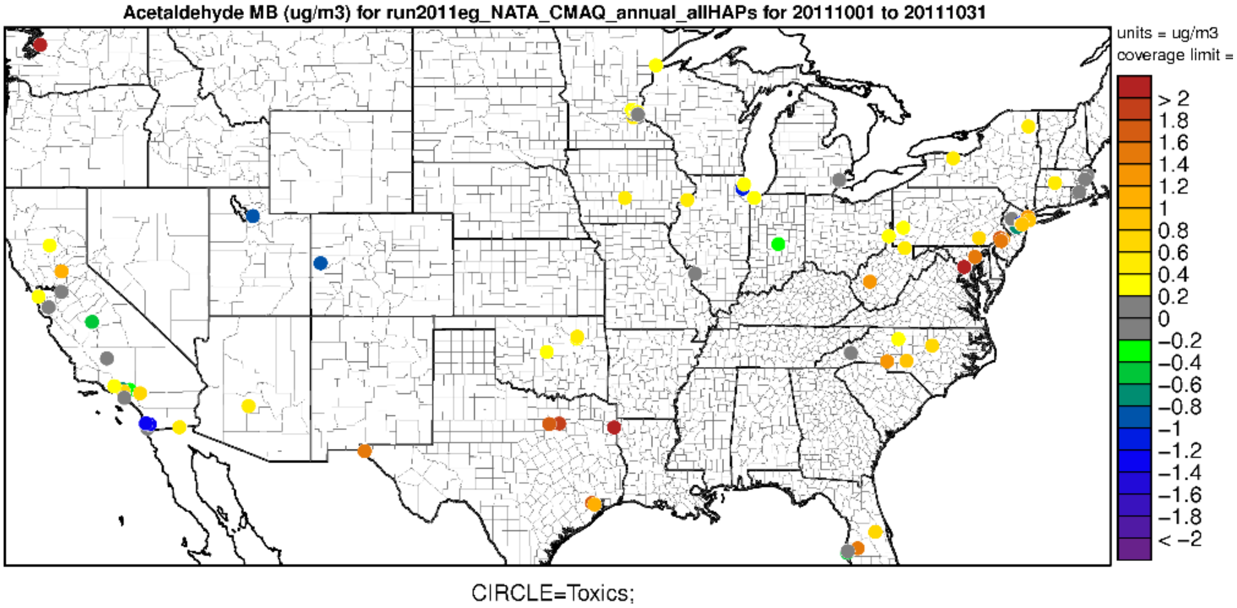


Exhibit 43. Acetaldehyde: 2011 Mean Error (%) at Monitoring Sites in the CMAQ Modeling Domain

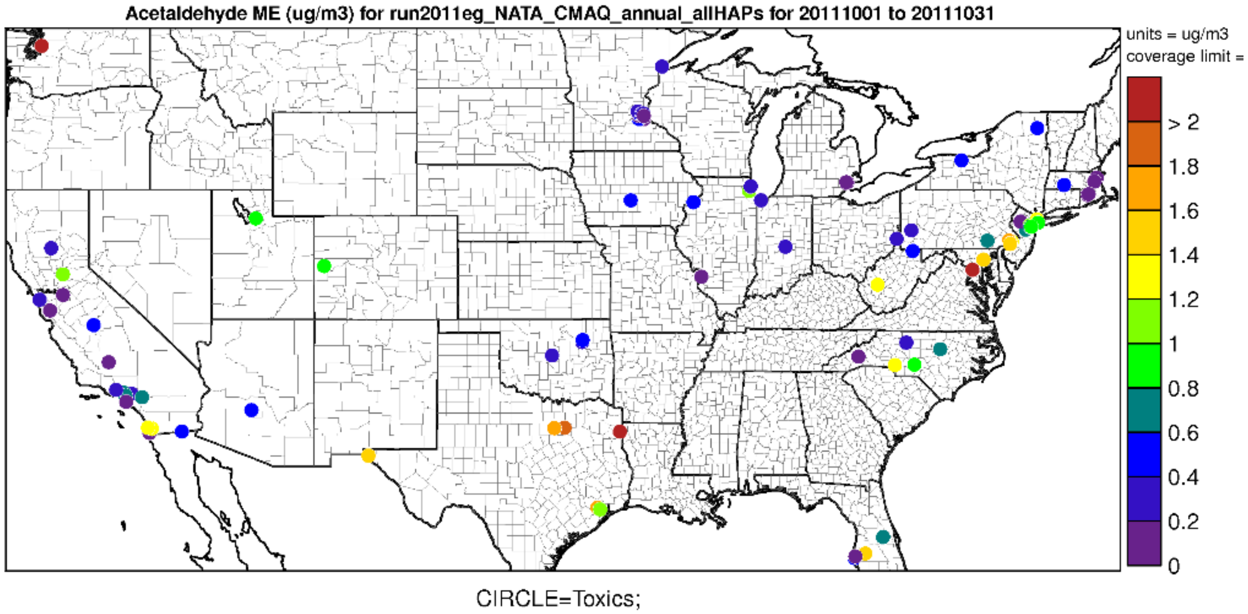


Exhibit 44. Acetaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the HEM-3 Modeling Domain

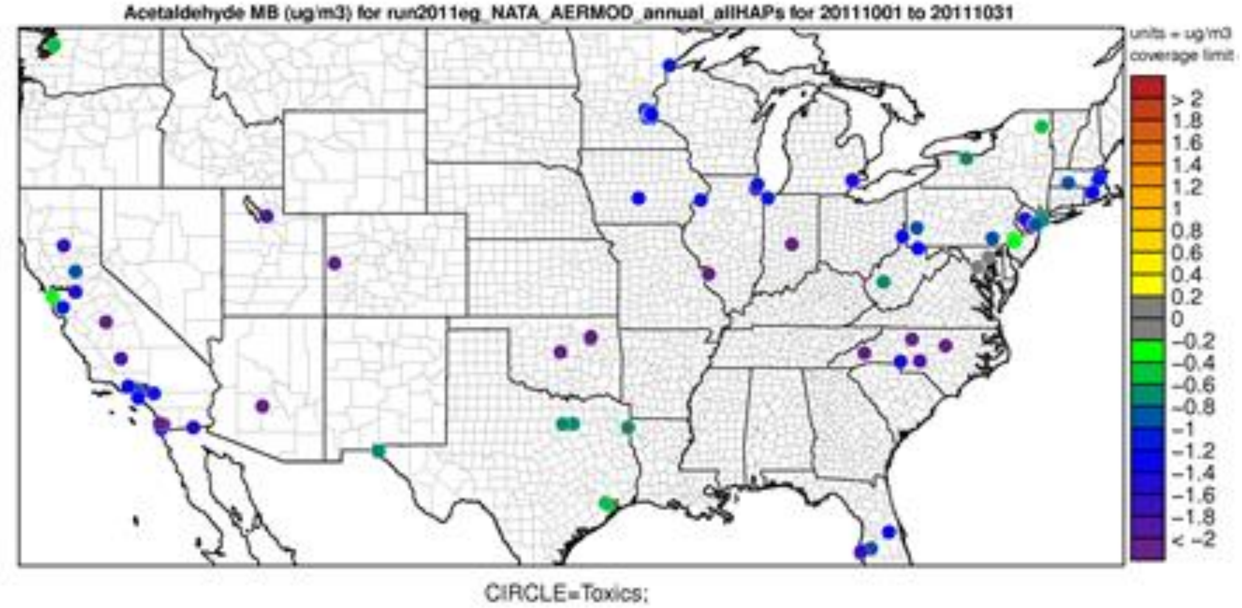


Exhibit 45. Acetaldehyde: 2011 Mean Error (%) at Monitoring Sites in the HEM-3 Modeling Domain

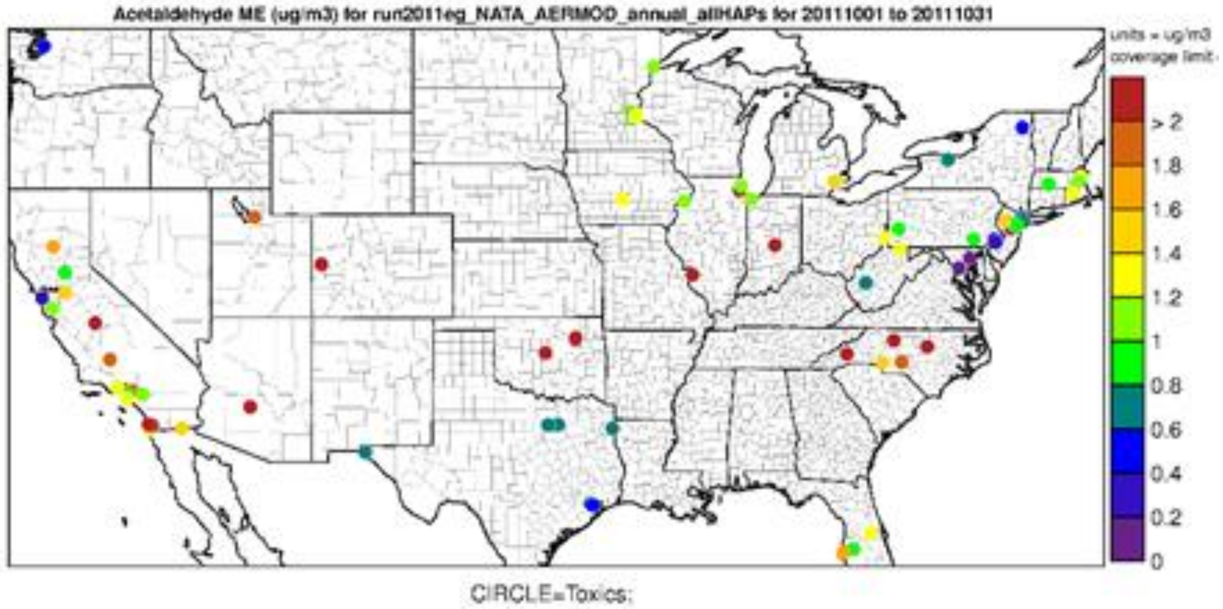
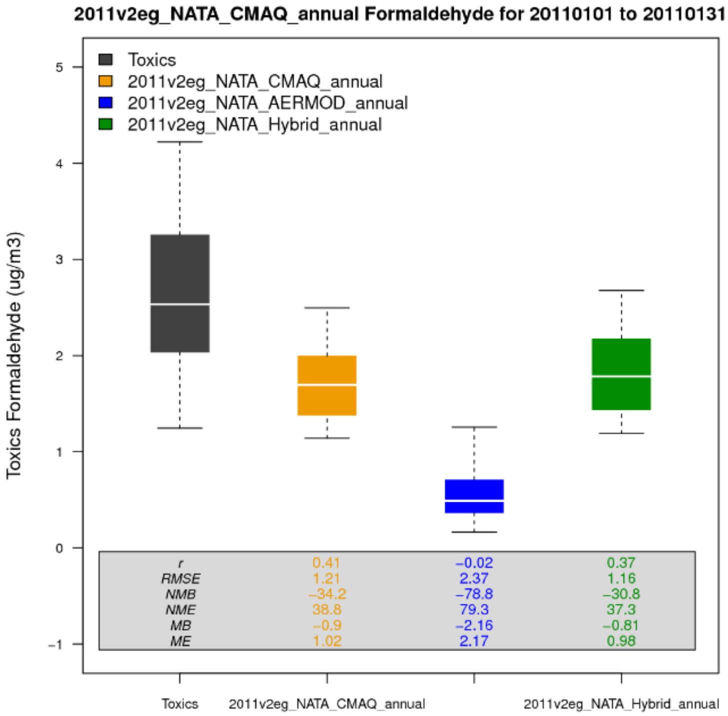


Exhibit 46. Formaldehyde: 2011 Boxplots of Observed and Modeled Concentrations (top) and Modeled-Observed Bias Difference in Concentrations (bottom) for the Hybrid, CMAQ, and HEM-3 Models



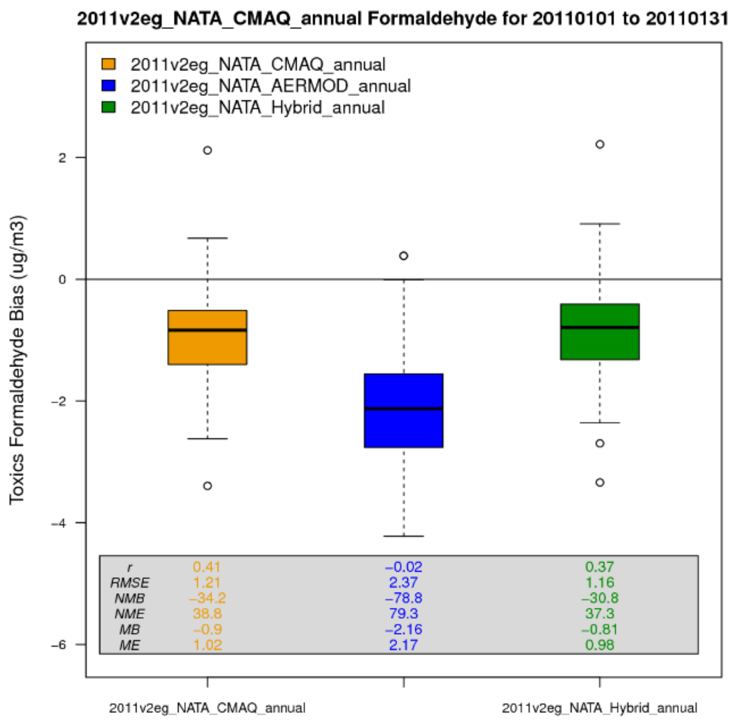


Exhibit 47. Formaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the Hybrid Modeling Domain

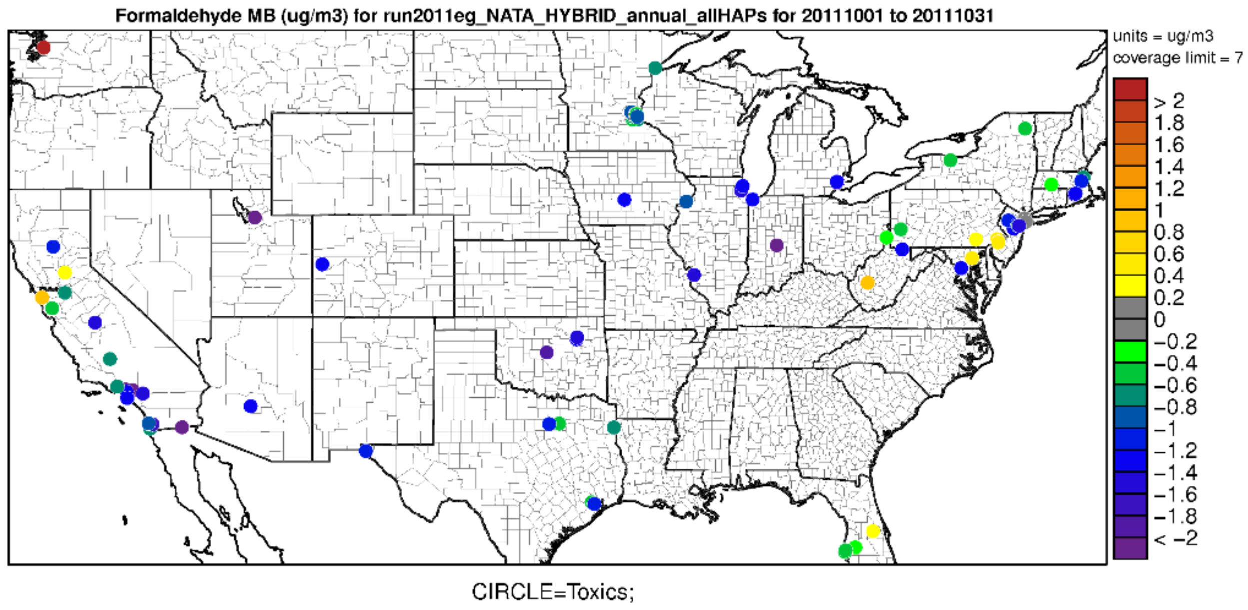


Exhibit 48. Formaldehyde: 2011 Mean Error (%) at Monitoring Sites in the Hybrid Modeling Domain

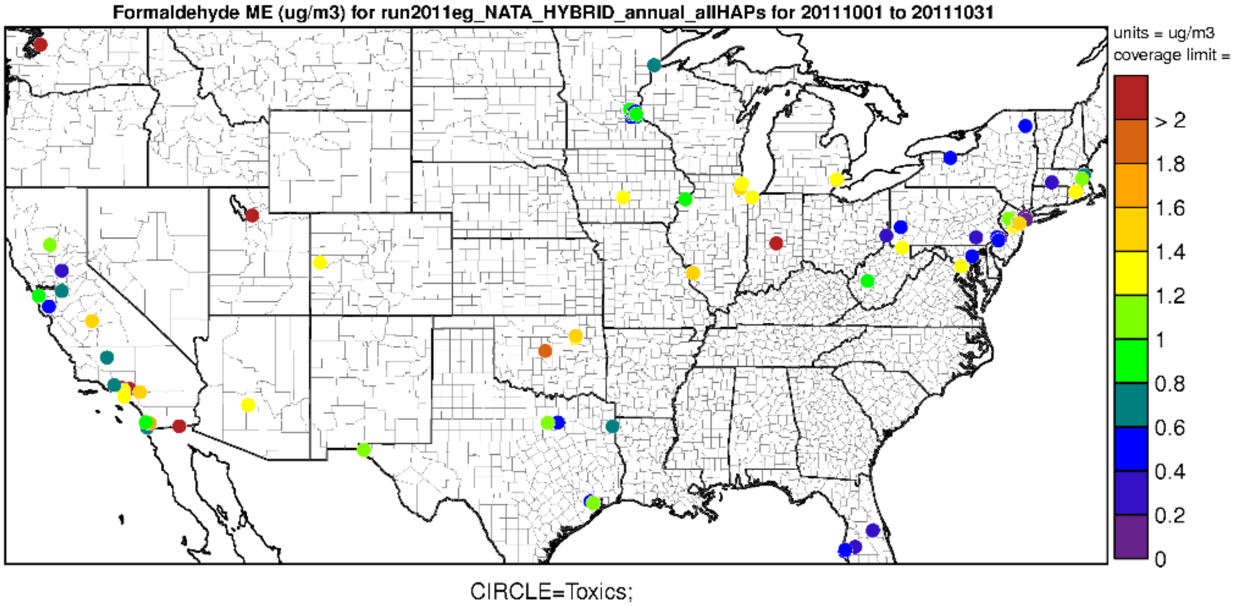


Exhibit 49. Formaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the CMAQ Modeling Domain

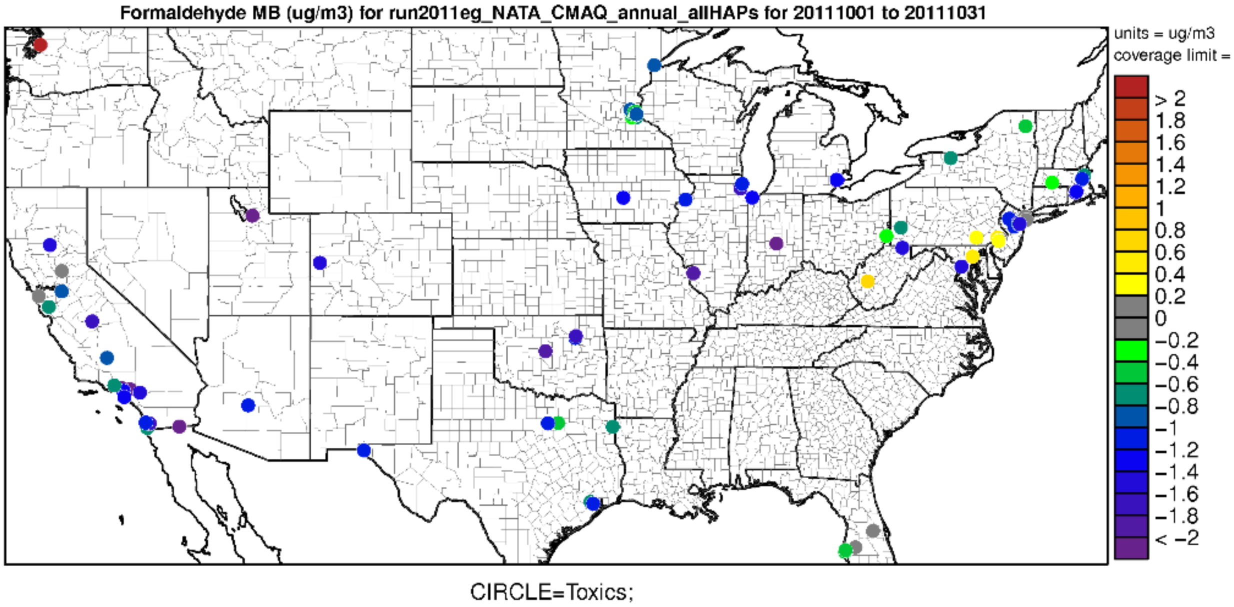


Exhibit 50. Formaldehyde: 2011 Mean Error (%) at Monitoring Sites in the CMAQ Modeling Domain

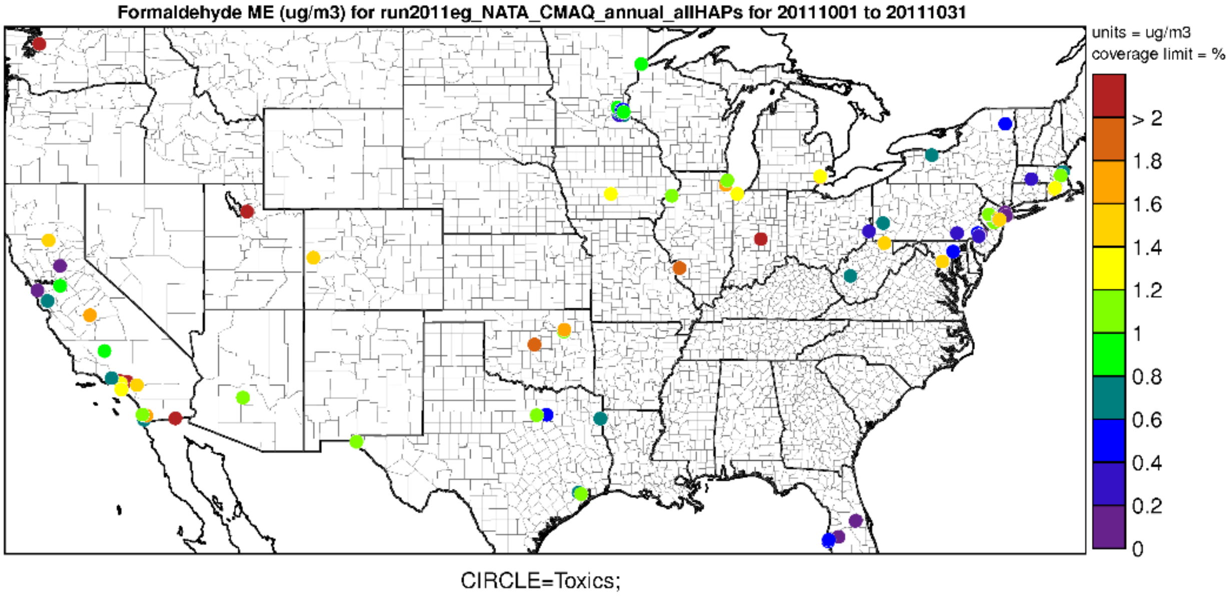


Exhibit 51. Formaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the HEM-3 Modeling Domain

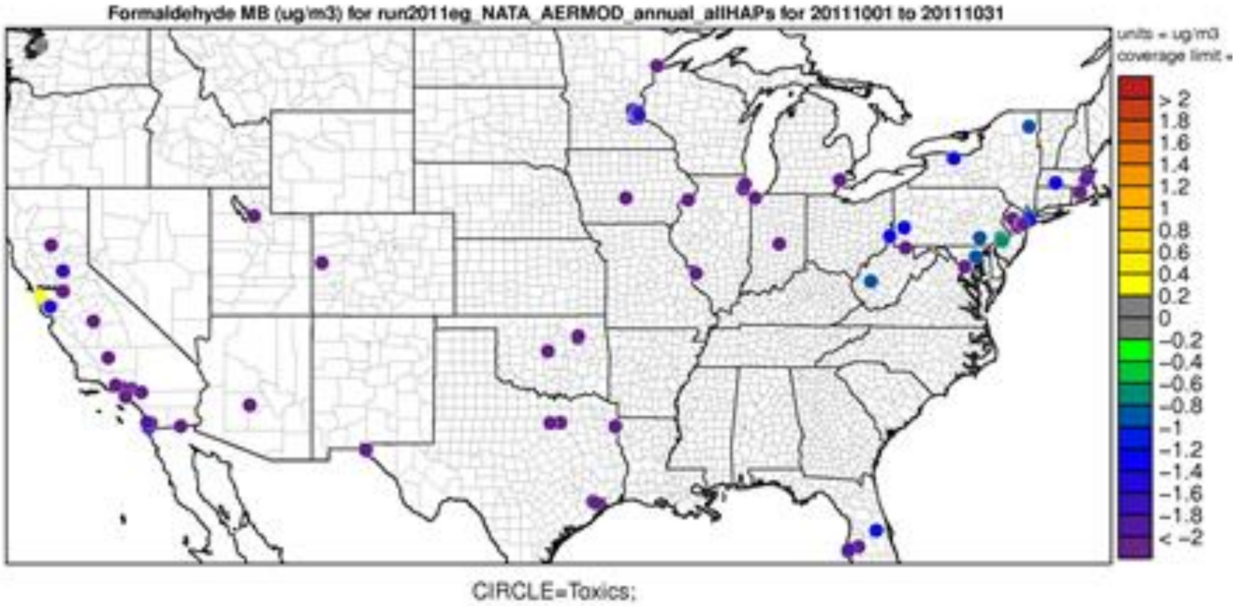


Exhibit 52. Formaldehyde: 2011 Mean Error (%) at Monitoring Sites in the HEM-3 Modeling Domain

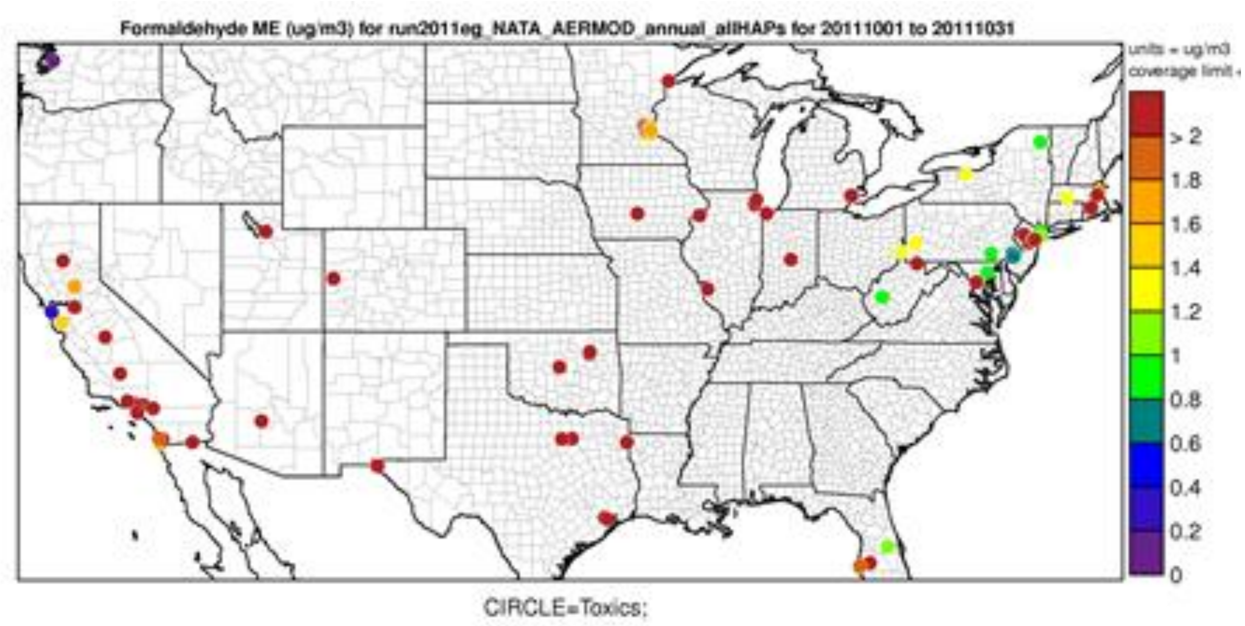


Exhibit 53. Benzene: 2011 Boxplots of Observed and Modeled Concentrations (top) and Modeled-Observed Bias Difference in Concentrations (bottom) for the Hybrid, CMAQ, and HEM-3 Models

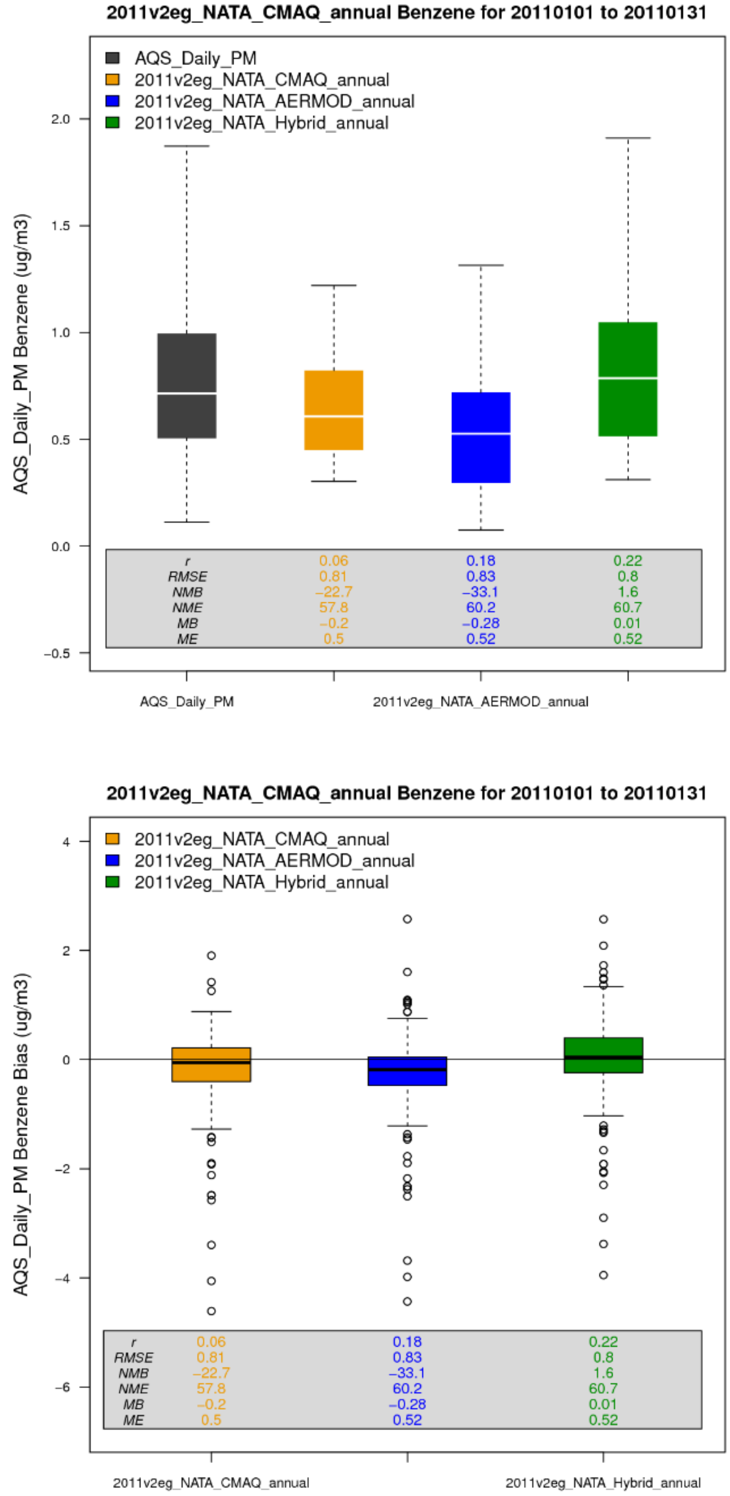


Exhibit 54. Benzene: 2011 Mean Bias (%) at Monitoring Sites in the Hybrid Modeling Domain

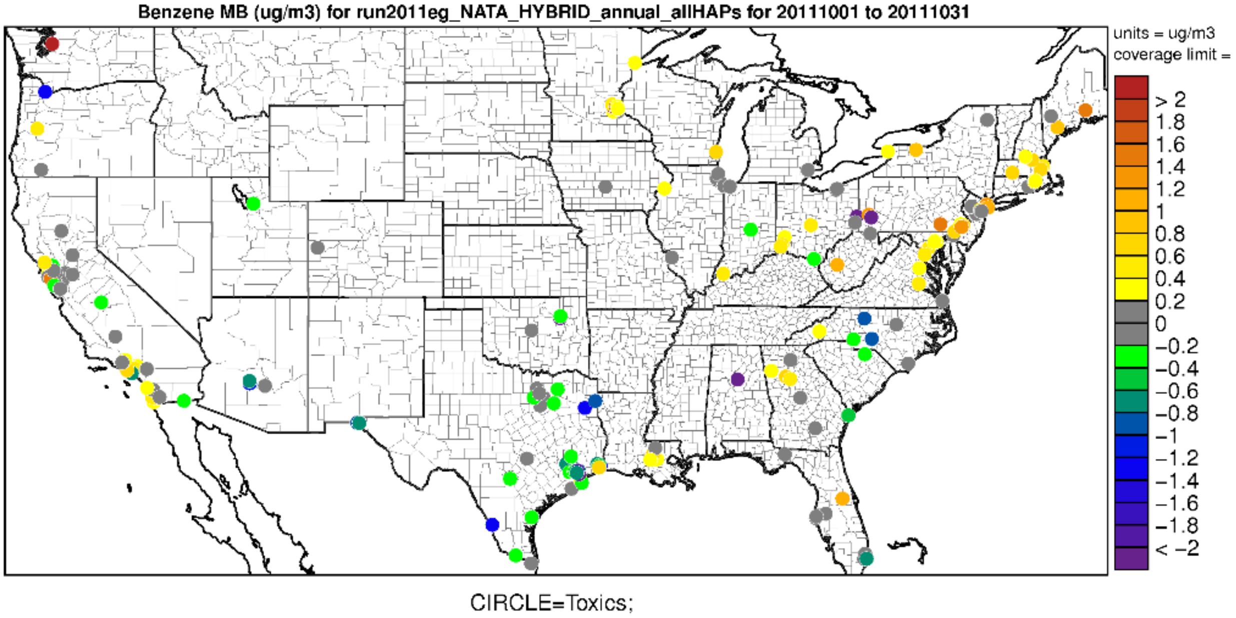


Exhibit 55. Benzene: 2011 Mean Error (%) at Monitoring Sites in the Hybrid Modeling Domain

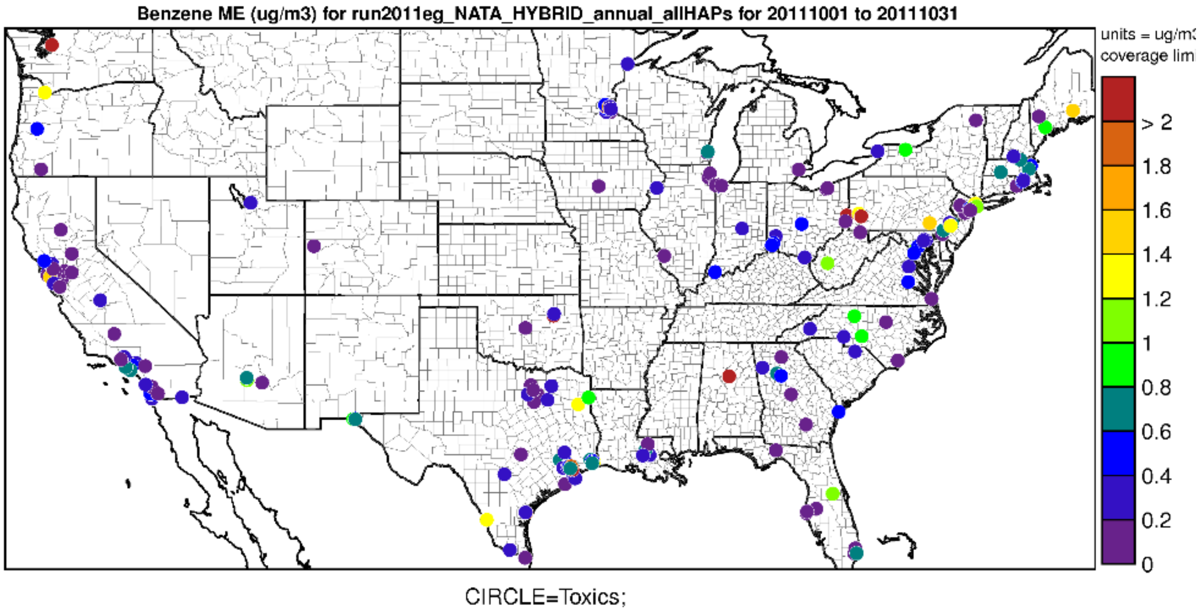


Exhibit 56. Benzene: 2011 Mean Bias (%) at Monitoring Sites in the CMAQ Modeling Domain

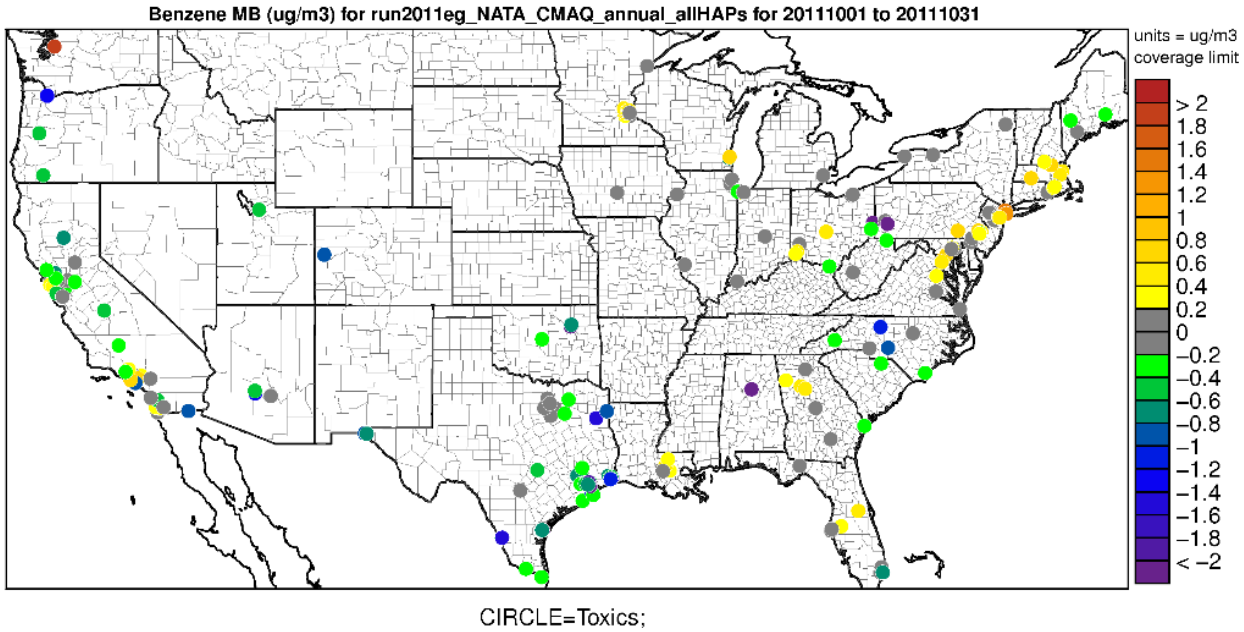


Exhibit 57. Benzene: 2011 Mean Error (%) at Monitoring Sites in the CMAQ Modeling Domain

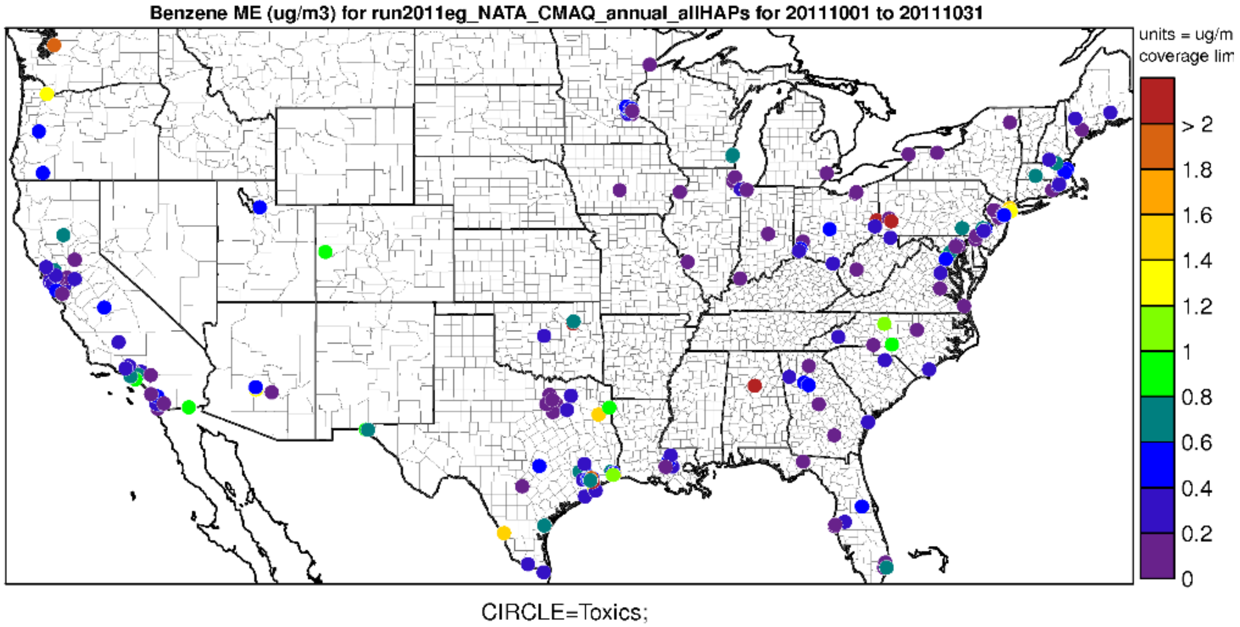


Exhibit 58. Benzene: 2011 Mean Bias (%) at Monitoring Sites in the HEM-3 Modeling Domain

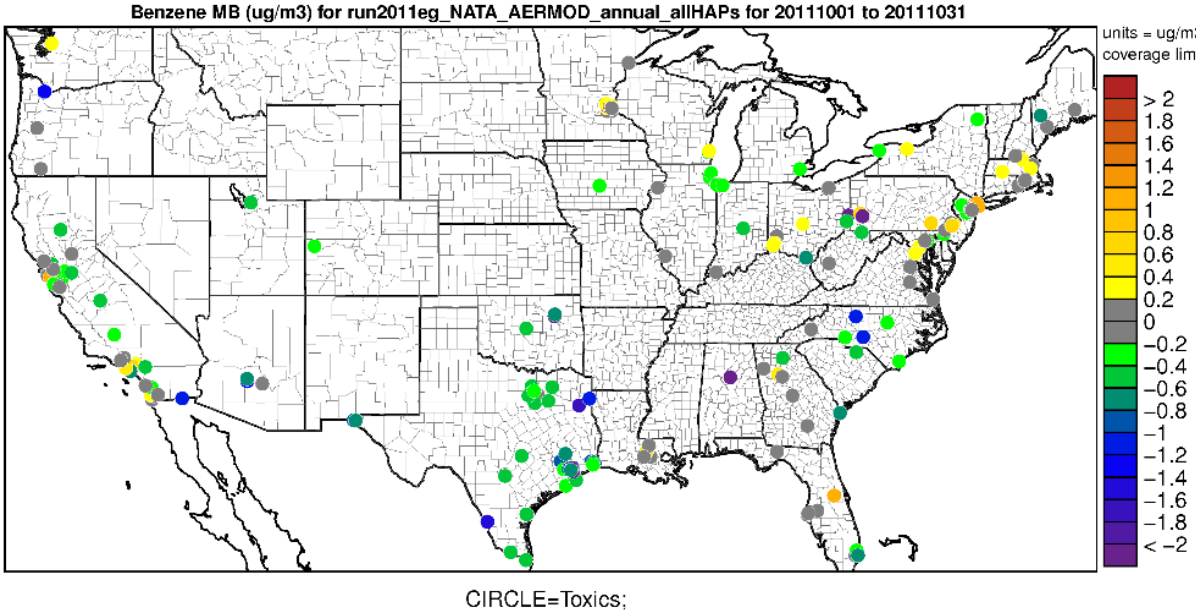
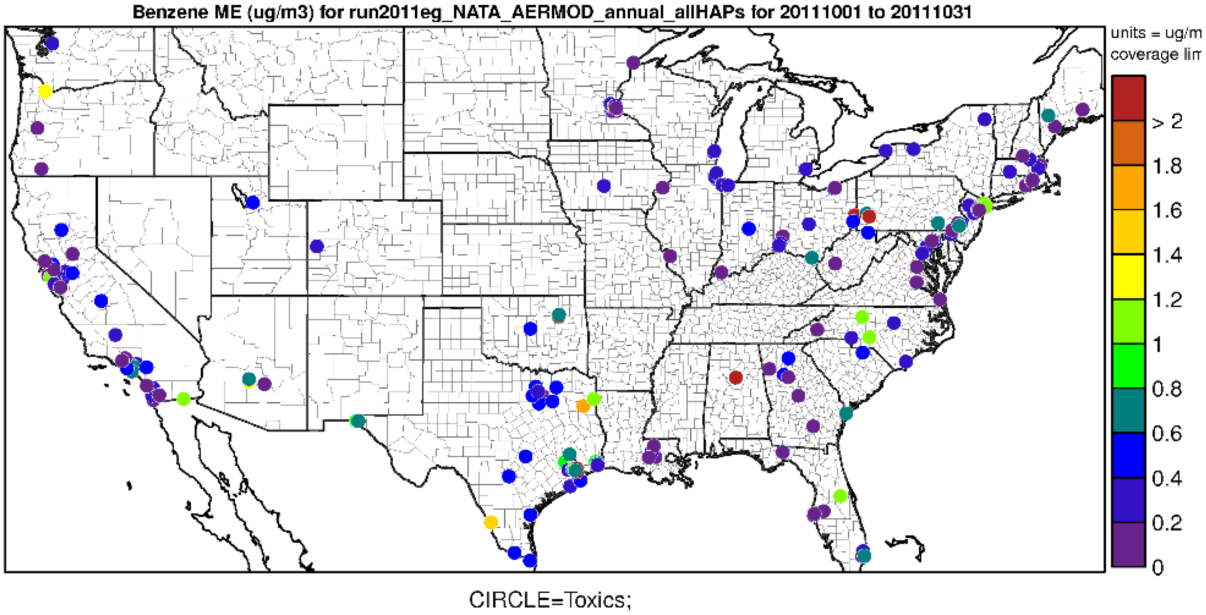


Exhibit 59. Benzene: 2011 Mean Error (%) at Monitoring Sites in the HEM-3 Modeling Domain



3.3.5 Non-hybrid Evaluation

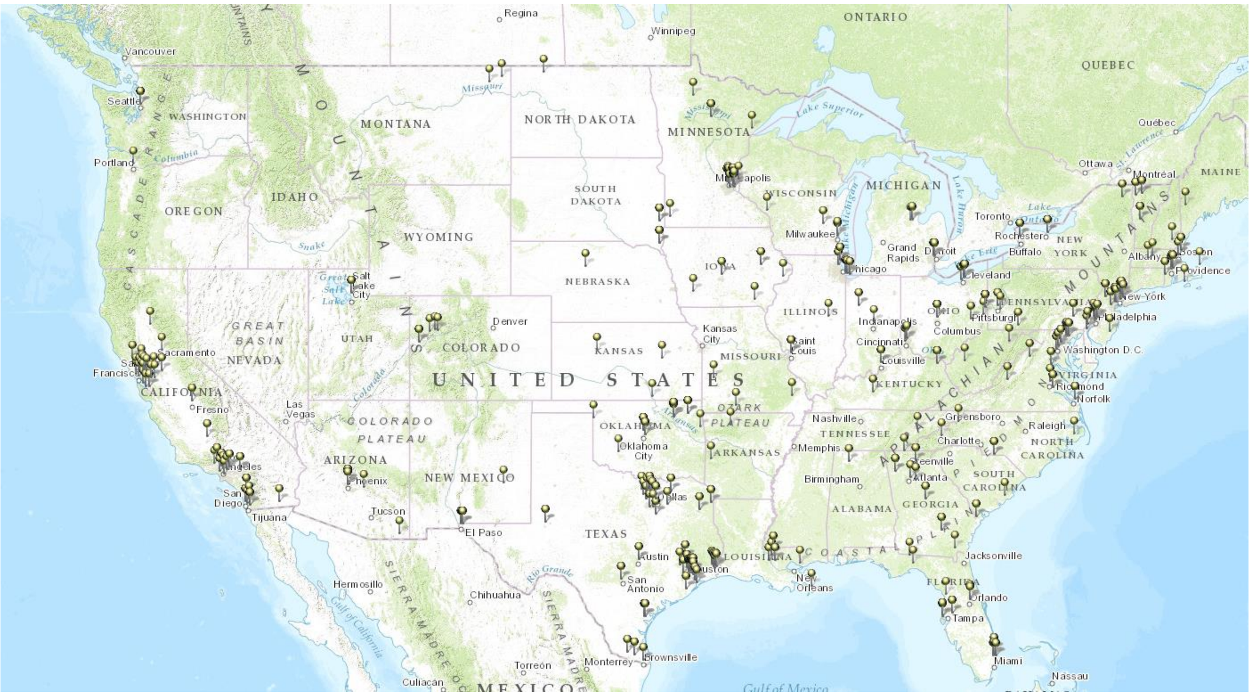
To estimate the ability of HEM-3 to replicate the 2011 observed ambient concentrations of air toxics, we conducted an annual operational model performance evaluation for air toxics used in the non-hybrid model calculation. Statistical assessments of modeled versus observed concentrations were paired in time and space and aggregated on an annual basis. Exhibit 60 contains a list of air toxics evaluated in the non-hybrid model performance evaluation and the number of paired sites (based on completeness criteria of observations, Section 3.3.2) used in the annual average. Exhibit 61 is a map of the 2011 monitoring locations for non-hybrid air toxics. Complete results from the non-hybrid evaluation are presented in Appendix F; paired values of the monitored and modeled data are provided in the spreadsheet "nonhybrid_polls-model_evaluation_paired.xlsx" in the SupplementalData folder. Data-completeness and measurement-uncertainty issues exist for: 1,1,1-trichloroethane, 1,1,2-trichloroethane , 1,2,4-trichlorobenzene, alpha-Chlorotoluene, hexachloro-1,3-butadiene, 1,1-dichloroethane , 1,1-dichloroethylene, chloroethane, and 3-chloropropene; therefore, these air toxics were not included in the non-hybrid evaluation. Also excluded were air toxics with no monitoring data available.

Exhibit 60. Non-hybrid Air Toxics Evaluated

Non-hybrid Air Toxic	Number of Paired Sites
4-Methyl-2-pentanone	35
Acetonitrile	36
Antimony (TSP) LC	11
Antimony PM_{10} LC	18
Bromomethane	26
Carbon disulfide	23
Carbon tetrachloride	163
Chloromethane	155
Cobalt (TSP) LC	12
Cobalt PM_{10} LC	15
Ethylbenzene	170

Non-hybrid Air Toxic	Number of Paired Sites
Hexane	125
Isopropylbenzene	23
Methyl tert-butyl ether	6
Propanal	68
Selenium (TSP) LC	11
Selenium PM_{10} LC	14
Selenium $PM_{2.5}$ LC	58
Styrene	71
Tribromomethane	5
Vinyl Acetate	25

Exhibit 61. 2011 Monitoring Locations for the Evaluation of Non-hybrid Air Toxics



3.4 Summary

- Forty of the more prevalent and higher-risk air toxics, accounting for 99 percent of the national human-health risk of all NATA air toxics, were modeled for air transport using a hybrid approach with CMAQ and HEM-3 (with AERMOD).
- Redundancies and double counting were minimized in the hybrid-modeling approach by anchoring air concentrations to those estimated by CMAQ.
- All other air toxics were modeled solely in HEM-3 (with AERMOD), except for background concentrations, which were estimated using monitoring and emissions data.
- Model evaluation generally shows better agreement for air toxics with more measurement sites and fewer measurement uncertainties, such as benzene.
- Characterization of model performance for NATA is significantly constrained given the multiple air toxic species assessed, limited air toxics observations and monitoring network, as well as multiple species across multiple spatial scales of interest. There are no accepted benchmarks for air toxics evaluation, in contrast to ozone model performance where error and bias for paired (monitoring site and hourly resolution) comparisons typically are within 35 and 15 percent, respectively (Russell and Dennis 2000; Seigneur and Dennis 2010). Such criteria are not realistic for air toxics due to characterization uncertainty in air toxics data.

This page intentionally left blank.

.

4 ESTIMATING EXPOSURES FOR POPULATIONS

Estimating inhalation ECs is a critical step in determining potential health risks because ambient concentrations do not take into account movements of individuals among geographic locations and microenvironments where pollutant concentrations can differ. Different individuals have different daily activities, spend different amounts of time engaged in those activities, and engage in those activities in different locations. Most activities occur in indoor environments (e.g., the home, workplace, school, and vehicles), where pollutant concentrations can differ from those in the outdoor environment. Therefore, the average concentration of a pollutant that people breathe can differ significantly from the ambient concentration at a fixed outdoor location.

Please refer to the version of this document accompanying the 2005 NATA (EPA 2011) for descriptions of exposure estimations for previous NATAs. This section contains a discussion of how ECs were estimated for the 2011 NATA. It begins with an overview of the hybrid approach used that included new exposure modeling for some pollutants and applications of exposure-to-ambient concentration ratios for the remaining NATA pollutants. This introduction is followed by a more detailed description of this approach, a summary of the user inputs and other data required, and an overview of the quality-assurance measures included in estimating exposures. Further details on the exposure calculations for the 2011 NATA can be found in Appendix G.

4.1 Estimating Exposure Concentrations

For the 2011 NATA, EPA used a hybrid of direct modeling and exposure factors to estimate inhalation ECs for NATA. This approach used census-tract-level ambient concentrations estimated with air quality models, as described in Section 3, and yielded census-tract-level exposure-concentration estimates that were used to determine potential health risks for NATA.

Direct exposure modeling used the EPA Hazardous Air Pollutant Exposure Model (HAPEM). HAPEM, described in detail in Section 4.2, is a screening-level exposure model that estimates inhalation ECs corresponding to estimated ambient-pollutant concentrations. EPA used version 7 of HAPEM (i.e., **HAPEM7**) for the 2011 NATA for a selected group of pollutants. For each modeled pollutant and NATA category (i.e., point, nonpoint, onroad mobile, and nonroad mobile), and for each census tract, EPA also calculated the ratio of EC to ambient concentration (i.e., an exposure factor). For the remaining pollutants not modeled with HAPEM, EPA selected one of the modeled pollutants as a surrogate, and sometimes the selected surrogate was specific to a NATA category. Per census tract, EPA multiplied the ambient concentration of the pollutant by the surrogate's exposure factor, resulting in estimated ECs. Section 4.4 contains additional description of this **exposure-factor approach**.

4.2 About HAPEM

Nearly two decades ago, EPA developed HAPEM for Mobile Sources (HAPEM-MS) to assess inhalation exposure to air toxics from highway mobile sources. This initial version of HAPEM used carbon monoxide as a tracer for highway mobile-source air toxic emissions. EPA has since updated and improved HAPEM to enable the prediction of inhalation ECs for a wide range of air toxics using either modeled ambient concentrations or measured data (without regard to source category), and the model no longer uses carbon monoxide as a tracer. More recent versions of HAPEM incorporate a range of enhancements, and, as a result, HAPEM version 4 and later versions can be used to predict annual average human-exposure levels on a nationwide basis at a spatial resolution as fine as the census-tract level (EPA 2002b, EPA 2005d, EPA 2007, EPA 2015c). The enhancements incorporated into recent versions of HAPEM facilitate its use for regional and national inhalation risk assessments such as NATA. Inhalation ECs for the 1996, 1999, and 2011 NATAs were estimated using HAPEM4, HAPEM5, and HAPEM7, respectively (EPA did not use HAPEM6 for NATA). Exhibit 62 outlines some key differences between these three versions. A complete history of HAPEM can be found in the User's Guide for HAPEM7 (EPA 2015c), the latest version of HAPEM available at the time this document was prepared.

Exhibit 62. Key Differences between Recent Versions of HAPEM

Characteristic	HAPEM4	HAPEM5	HAPEM7
Data source for population demographics	1990 U.S. Census	2000 U.S. Census	2010 U.S. Census
Characterization of microenvironmental factors	Point estimates	Probability distributions	Same as HAPEM5
Method for creation of annual-average activity patterns from daily activity-pattern data	Resampling of daily diaries for each of 365 days without accounting for autocorrelation	Sampling a limited number of daily diaries to represent an individual's range of activities, accounting for autocorrelation	Same as HAPEM5, except now includes commuter-status criterion
Interpretation of exposure-concentration range for a given cohort/tract combination	Uncertainty for the average annual EC for the cohort/tract combination	Variability of annual ECs across cohort/tract members	Same as HAPEM5, except now includes adjustments based on proximity to roadway

HAPEM uses a general approach of tracking representative individuals of specified demographic groups as they move among indoor and outdoor microenvironments and among geographic locations. As described in the following section, personal-activity and commuting data specific to a hypothetical individual's demographic groups are used to determine the census tracts containing residential and work locations and the microenvironments within each tract. Empirically based factors reflecting the relationship between ECs within each microenvironment and the outdoor (ambient) air concentrations at that location are selected by the model through a stochastic sampling process to estimate ECs.

To estimate long-term ECs for a hypothetical individual, the pollutant concentrations in each microenvironment visited are first combined into a daily-average concentration. The daily averages are then combined with proper weighting for season and day type to calculate a long-term average. Finally, the long-term averages are stratified by demographic group and census tract to create a distribution of ECs for each stratum. The median of each distribution represents the best estimate of exposure for a "typical" person of that demographic group in that census tract. In this case, "typical" does not refer to a specific individual in the population or even the average over a group of individuals. Rather, this person is a hypothetical individual residing at the centroid of a census tract and engaging in a range of activities (both indoor and outdoor) representative of those in which individuals of that demographic group in that census tract might engage. Additional technical information on HAPEM can be found in the User's Guide for HAPEM7 (EPA 2015c).

4.3 HAPEM Inputs and Application

HAPEM requires four primary types of information to estimate ECs: (1) ambient concentrations of air toxics, (2) population data from the U.S. Census Bureau, (3) population-activity data, and (4) microenvironmental data. The subsections below contain additional discussion on these inputs, accompanied by descriptions of the data used for NATA and related information on how EPA configured the model and applied it to conduct direct exposure modeling.

4.3.1 Data on Ambient Air Concentrations

HAPEM is typically applied using annual-average, diurnally distributed ambient air concentrations. Input concentrations can be monitoring data or concentrations estimated using a dispersion model or other air quality model.

For the 2011 NATA, EPA estimated annual-average ambient concentrations for each census tract using a hybrid CMAQ-HEM-3 approach discussed in Section 3. EPA stratified the air quality outputs for a selected group of pollutants by one or more of the four principal NATA categories (i.e., point, nonpoint, onroad mobile, and nonroad mobile), and EPA used those results as surrogates for the remaining pollutants not modeled in CMAQ-HEM-3. Thus, exposure-model results generated for NATA can be summarized for each principal NATA category or any combination of those categories.

4.3.2 Population Demographic Data

HAPEM divides the exposed population into cohorts such that each person in the population is assigned to one and only one cohort, and all the cohorts combined encompass the entire population. A cohort is defined as a group of people whose exposure is expected to differ from exposures of other cohorts due to certain characteristics shared by the people within that cohort. In HAPEM7 used for the 2011 NATA, cohorts were defined using residential census tract and age so that the population in each census tract was divided into six age groups 0–1, 2–4, 5–15, 16–17, 18–64, and ≥ 65 years of age. These groups were developed using demographic data derived from the 2010 U.S. Census. EPA aggregated the predicted inhalation ECs across cohorts to estimate ECs for the general population.

4.3.3 Data on Population Activity

HAPEM draws on two types of data to define activities for the modeled population: activity-pattern data (specifying the frequency, location, and duration of daily activities) and commuting-pattern data (specifying the work tracts for people living in each home tract). HAPEM uses these data in coordination to place a hypothetical individual who commutes to work either in the home tract or the work tract and in a specific microenvironment at each 3-hour time step (the time step used for NATA). The microenvironment assignments and locations derived from these data are then used to calculate ECs, as explained in the next section. EPA discusses these two types of data in more detail in the following paragraphs.

Data on **human activity patterns** are used to determine the frequency and duration of exposure within various microenvironments such as indoors at home, in-vehicle, and outdoors. Activity-pattern data are taken from demographic surveys of individuals' daily activities that specify the sequence, duration, and locations of those activities. The default source of activity-pattern data used by HAPEM and for NATA is EPA's Consolidated Human Activity Database (CHAD; EPA 2015a). To develop the version of CHAD used in the 2011 NATA (i.e., version June 2014), data from 21 individual U.S. studies of human activities were combined into one comprehensive data system that contains over 45,000 person-days of activity-pattern records (previous NATAs used previous versions of CHAD containing of fewer studies and person-days of data). Because of limitations of the study designs of the surveys from which it is derived, CHAD might not be representative of all demographic groups, particularly ethnic minorities and low-income populations. Another limitation of the activity-pattern data in CHAD is that most are for individuals over a one- or two-day period only. Extrapolation of these short-term records to the annual activity patterns required for assessments of air toxics exposure introduces some uncertainty into the analysis.

The algorithms in HAPEM address this extrapolation uncertainty by implementing a stochastic process to create simulated long-term (multi-day) activity patterns from daily activity-pattern data that account for day-to-day autocorrelation. These algorithms create annual-average activity patterns from daily activity-pattern data to represent the variability more fully among individuals within a cohort-tract combination. For each day type and demographic group, daily-activity diaries were divided into three groups based on similarity using a cluster analysis. To simulate the activities of an individual, one diary was selected from each group for each day type, resulting in nine diaries in total. Then, for each day type, the sequence of the selected diaries was determined according to the probability of transition from one cluster group to another, as determined by analyses of the CHAD data. The simulation was repeated 30 times, resulting in a set of 30 estimates of annual ECs for each demographic group in each census tract. Use of a limited number of diaries and the transition probabilities is a way to account for day-to-day autocorrelation of activities for an individual, so each exposure-concentration estimate represents an estimate for an individual rather than an average for the group. Therefore, with this approach, the range represents the variability of ECs across the group. These algorithms were used beginning with HAPEM5 (i.e., beginning with the 1999 NATA).

Commuting-pattern data, the second type of population activity data used in HAPEM, are derived for each cohort from a U.S. Census database containing information on tract-to-tract commuting patterns. These data specify the number of residents in each tract that work in that tract and every other census tract (i.e., the population associated with each home-tract/work-tract pair) and the distance between the centroids of the two tracts. An important limitation is that the commuting-pattern data included in HAPEM do not account for the movement of school-age children who travel (or commute) to a school located outside of their home tract.

4.3.4 Microenvironmental Data

A microenvironment is a three-dimensional space in which human contact with an environmental pollutant occurs. In HAPEM, this space is treated as a well-characterized, relatively homogenous location with respect to pollutant concentrations for a specified period. The inhalation exposure estimate is determined by the sequence of microenvironments visited by the individual. The concentration in each microenvironment is estimated by using the three microenvironmental factors listed below to adjust the ambient-concentration estimate for the census tract where it is located:

- a **penetration factor** that is an estimate of the ratio of the microenvironmental concentration to the concurrent outdoor concentration in the immediate vicinity of the microenvironment; penetration factors are pollutant-specific estimates that are derived from reported measurement studies;
- a **proximity factor** that is an estimate of the ratio of the outdoor concentration in the immediate vicinity of the microenvironment to the outdoor concentration represented by the ambient air concentration input to the model; and
- an **additive factor** that accounts for emission sources within or near a particular microenvironment, such as indoor emission sources. As noted below, the additive factor is not used for NATA.

The relationship between the estimated ECs, the input ambient concentrations, and these three factors is demonstrated by the equation below.

$$C_{(i,k,t)} = CONC_{(i,t)} \times PEN_k \times PROX_k + ADD_k$$

Where:

$C_{(i,k,t)}$ = EC predicted within census tract *i* and microenvironment *k* for time step *t*, in units of $\mu g/m^3$

$CONC_{(i,t)}$ = ambient concentration for census tract *i* for time step *t*, in units of $\mu g/m^3$

PEN_k = penetration factor for microenvironment *k*

$PROX_k$ = proximity factor for microenvironment *k*

ADD_k = additive factor accounting for sources within microenvironment *k*, in units of $\mu g/m^3$

Stochastic processes can be used to select work tracts, ambient air concentrations, and microenvironmental factors. This important feature allows exposures to be characterized with probability distributions rather than point estimates, which more accurately reflect the variability of these components and simulate some of the variability found in measurement studies.

In HAPEM, the characteristics of each microenvironment are used to assign each microenvironment to one of three groups: indoors, outdoors, and in-vehicle. The 2011 NATA used the 18 microenvironments shown in Exhibit 63. The microenvironments in the indoor group were further classified as associated with either residence or other buildings, while those in the outdoor group were categorized as either near-road or away-from-road. Each group consists of microenvironments expected to have similar penetration factors, thus allowing microenvironmental factors developed for one microenvironment to be applied to other microenvironments in the same group. Within each census tract, HAPEM uses estimates of the number of people living within each of three distance-from-road bins to stochastically vary the proximity factor based on distance-from-road (i.e., proximity factors are higher for microenvironments near major roadways, lower for microenvironments relatively far from major roadways). The additive factor (ADD_k) in the expression for EC, above, was set to zero for NATA because indoor-source data are currently incomplete (recall that NATA covers only pollutants derived from outdoor sources).

An important consideration is that data to support quantitative microenvironmental factors are not well developed for many of the air-toxic compounds and for most of the microenvironments, which introduces uncertainty into the analysis of exposures. Section 7 contains a discussion on uncertainty and variability with regard to this and other issues for NATA.

Exhibit 63. Microenvironments Used in the HAPEM Modeling for the 2011 NATA

Indoors	Outdoors	In Vehicle
Residence	*Near-road*	Car/Truck
Residential	Motorcycle/Bicycle	Public Transit
Other Building	Outdoors, Near Roadway	
Air Travel	Outdoors, Parking Garage	
Bar/Restaurant	Outdoors, Service Station	
Hospital	Residential Garage	
Office	Waiting Outdoors for Public Transit	
Public Access	*Away-from-road*	
School	Ferryboat	
Waiting Inside for Public Transit	Outdoors, Other	

4.4 Exposure Factors

HAPEM exposure modeling for NATA requires substantial time and resources for data collection and processing, computing, and model processing. Due to these requirements, we conducted HAPEM modeling for the 2011 NATA only for selected pollutants, which we present below along with how we used them to estimate ECs for the remaining NATA pollutants.

- **Coke oven emissions** (emitted by point sources and present in ambient air as either particulates or gases) and **DPM** (modeled as particulates from nonpoint and mobile sources) were special cases that EPA modeled as themselves in HAPEM and not used as surrogates for any other pollutants not modeled in HAPEM.
- **Benzene** and **1,3-butadiene** are gas-phase pollutants emitted by many processes (and all four principal NATA categories) in nearly all U.S. locations. EPA selected benzene as the surrogate for all other gas-phase pollutants not modeled in HAPEM (EPA considers benzene modeling in NATA to be more reliable than 1,3-butadiene modeling).
- Unspeciated, generic PAHs ("**PAH, total**"), which are pollutants that can be present in either gas phase or particulate phase in ambient air, are emitted by all four principal NATA categories and from a wide variety of processes. EPA selected "PAH, total" as the surrogate for all other mixed-phase pollutants not modeled by HAPEM.
- **Chromium (VI)** is a highly toxic particulate-phase pollutant emitted by all four principal NATA categories, and EPA selected it as the surrogate for all other particulate pollutants not modeled in HAPEM and emitted by point or nonpoint sources.
- EPA selected **nickel**, a particulate-phase pollutant emitted by a variety of processes spread across the United States, as the surrogate for all other particulate pollutants not modeled in HAPEM and emitted by mobile sources.

Exhibit G-1 in Appendix G to this document contains the overall average exposure-to-ambient concentration ratios (i.e., exposure factors) calculated from HAPEM and air quality outputs for each pollutant. EPA presents this factor for each of the four principal NATA categories (i.e., point, nonpoint, onroad mobile, and nonroad mobile). EPA used factors specific to each census tract, so these averages are for summary/informational purposes only. Overall, the HAPEM exposure predictions are lower than the corresponding predicted air quality values. This reduction likely results from the inability of many pollutants to penetrate efficiently into an indoor environment. (Recall that indoor sources of air toxics have not been included in any versions of NATA completed to date).

4.5 Quality Assurance in Exposure Modeling

A model-performance evaluation can provide valuable information regarding model uncertainty when using computer-simulation models of human exposures to pollutants, and a well-conducted evaluation can substantially increase confidence in model results for a given application or use. One type of performance evaluation is the use of measurements and environmental data as a benchmark for comparison of modeling estimates. EPA has worked with the Mickey Leland Center (NUATRC 2011) on past assessments to help identify new and independent sources of personal-monitoring data for use in comparison with the NATA results.

Extensive peer review involving independent scientific and technical advice from scientists, engineers, and economists can be another valuable component of a model evaluation. In July 2000, HAPEM4 underwent external peer review by technical experts for both the microenvironmental factors used in the model and the overall application of the model for NATA. A discussion of several of the issues addressed by these reviews is included in Appendix A of the report for the 1996 NATA presented to EPA's Science Advisory Board for review (EPA 2001b). In 2001, EPA's Scientific Advisory Board reviewed the application of HAPEM4 as part of the 1996 NATA review (EPA 2001a). Although several limitations were identified in the current methodology, HAPEM4 was acknowledged as an appropriate tool to help better understand the relationship of human exposures to ambient-concentration levels.

4.6 Summary

- Estimating inhalation ECs is a critical step in determining potential health risks because ambient concentrations do not account for movements of individuals among geographic locations and microenvironments where pollutant concentrations can differ.
- We estimated inhalation ECs for each census tract for the 2011 NATA using direct exposure modeling with HAPEM7 for some pollutants and exposure-factor surrogates for the remaining pollutants.
- These tract-level ECs can be used to determine potential health risks.

This page intentionally left blank.

5 CHARACTERIZING EFFECTS OF AIR TOXICS

Exposure to air toxics is associated with increased incidence of cancer and a variety of adverse noncancer health effects. The type and severity of effects depends on several factors, including the identity and nature of the chemical to which an individual is exposed, the magnitude and duration of exposure, and the unique behaviors and sensitivities of exposed individuals. The process of identifying and quantifying the adverse health effects associated with exposure to a chemical is accomplished with EPA risk assessment methods by way of a toxicity assessment. As indicated in Exhibit 2 of this document and described in more detail in Volume 1 of EPA's ATRA Reference Library (EPA 2004a), two processes constitute toxicity assessment: hazard identification (during which the specific adverse effects are identified that can be causally linked with exposure to a given chemical) and dose-response assessment (which characterizes the quantitative relationship between chemical dose or concentration and adverse effects, that is, the hazard(s) identified in the first step).[10] Ultimately, the results of the toxicity assessment, referred to in this document as "toxicity values," are used in conjunction with exposure estimates to characterize the health risks for exposed populations as described in Section 6. Although the toxicity assessment is integral and important to the overall air toxics risk assessment, it is usually accomplished prior to the risk assessment. We have completed this toxicity assessment for many air toxics and have made available the resulting **toxicity information** and **dose-response values**, which have undergone **extensive peer review**.

This section contains explanations of how toxicity assessments are used in the NATA risk assessment process. Specifically, the sections that follow provide an overview of the cancer and noncancer toxicity values used in NATA and the primary sources of these values. Several adjustments and assumptions to toxicity values that are specific to the NATA risk assessment process are also described.

5.1 Toxicity Values and Their Use in NATA

The toxicity values used for NATA are quantitative expressions used to estimate the likelihood of adverse health effects given an estimated level and duration of exposure. These toxicity values are based on the results of dose-response assessments, which estimate the relationship between the dose and the frequency or prevalence of a response in a population or the probability of a response in any individual. Because **NATA is focused on long-term exposures**, the toxicity values used in NATA are based on the results of chronic dose-response studies when such data are available. Chronic dose-response assessments can be used to help evaluate the specific 70-year-average (i.e., "lifetime") ECs associated with cancer prevalence rates, or, for noncancer effects, the concentrations at which noncancer adverse health effects might occur given exposure over an extended period of time (possibly a lifetime, but the time frame also can be shorter).

The toxicity values that are combined with ECs to conduct the risk characterization in NATA are based on the results of quantitative dose-response assessments. The actual values used, however, are not strictly considered dose-response or concentration-response values. To estimate cancer risks in NATA, the results of cancer dose-response assessments for a given chemical were converted to a URE that incorporates

---

[10] The phrase "dose-response" is used generally throughout this document to refer to the relationship between a level of a chemical and a physical response. The values EPA uses for inhalation, however, are derived for exposure concentration, although with consideration of dose. Consideration of the relationship between exposure concentration, dose, and dosimetry (how the body handles a chemical once it is inhaled) is inherent in the derivation of values. The term "toxicity values" is used here to refer to the RfCs and UREs used in inhalation risk assessment.

certain exposure assumptions. This value can be multiplied by the 70-year-average EC to obtain a lifetime cancer risk estimate for each individual. To evaluate the potential for noncancer adverse health effects, chronic dose-response data were used to estimate a threshold that is the EC in air at which adverse health effects are assumed to be unlikely (i.e., the RfC). These two types of values are described in more detail in the following section.

The toxicity values used in NATA are consistent with those OAQPS has compiled for chronic inhalation exposures to air toxics. The full set of toxicity values (and the sources of the values) used for the 2011 NATA are found in Appendix H. Sources of chronic dose-response assessments used for the 2011 NATA were prioritized according to OAQPS risk assessment guidelines and level of peer review, as discussed below.

5.2 Types of Toxicity Values

Each toxicity value used in NATA is best described as an estimate within a range of possible values appropriate for screening-level risk assessments. Of importance to note is that the uncertainty in the dose-response assessments and toxicity values that NATA relies on is to some extent one-sided, providing a **conservative (health-protective) estimate of risk**. The "true" cancer risk and potential for adverse noncancer impacts are believed to be lower than those estimated in this assessment, although the possibility remains that they could be greater. Uncertainty in the derivation of the dose-response values and in other aspects of the NATA process is discussed in Section 7.

5.2.1 Cancer URE

The **URE** is the upper-bound excess lifetime cancer risk estimated to result from continuous exposure to an agent at a concentration of 1 microgram per cubic meter (µg/m^3) in air. UREs are considered upper-bound estimates, meaning they represent a plausible upper limit to the true value. The true risk is likely to be less, but could be greater.

A cancer dose-response curve is used to demonstrate the quantitative relationship between dose and the likelihood of contracting cancer. If the dose-response relationship is linear, the cancer response is assumed to increase proportionally with the dose (which might be expressed as an EC, an absorbed internal dose, a dose to a specific organ or tissue, or other measure). We have proposed that linear extrapolation of carcinogenic risk in the low-dose region of the curve is a **reasonable approach for estimating risk at relatively low exposures**, such as those typically experienced by the general population for air toxics (i.e., the true value of the risk is unknown, and could be as low as zero). An upper-bound lifetime cancer risk represents a plausible upper limit to the true probability that an individual will contract cancer as a result of exposure over a 70-year lifetime to a given hazard (e.g., exposure to an air toxic).

For an inhalation risk assessment (and for NATA), a URE can be used to calculate the estimated cancer risk from inhalation ECs. A URE is calculated by using dose-response information for a chemical and developing a factor in the appropriate units that can be combined directly with ECs in air to estimate individual cancer risks, given certain assumptions regarding the exposure conditions. Specifically, the URE represents the upper-bound of the excess cancer risk estimated to result from continuous exposure to a concentration of 1 µg of a substance per m^3 of air, over a 70-year lifetime and assuming a daily inhalation rate of about 20 m^3/day. The risk value is derived from the slope of the dose-response curve as estimated using a linearized multistage statistical model in the low-dose portion of the curve. The interpretation of the URE is as follows: If the URE is 1.5×10^{-6} µg/m^3, no more than 1.5 excess tumors would develop per 1,000,000 people if they were exposed daily for a lifetime to a concentration of 1 µg/m^3. To the extent that true dose-response relationships for some air toxics compounds are not strictly linear, this assumption could result in overestimates of cancer risk. The upper bound is not a true statistical confidence limit because the URE reflects unquantifiable assumptions about effects at low doses. Thus, although **the actual carcinogenic risk is likely to be lower than what is reflected in the URE**, it also might be higher.

The URE provides an estimate of toxic potency of a chemical. EPA's weight-of-evidence (WOE) descriptors provide estimates of the level of certainty regarding a chemical's carcinogenic potential. We evaluate three broad categories of toxicological data to make a WOE determination: (1) human data (primarily epidemiological); (2) animal data (results of long-term experimental animal bioassays); and (3) supporting data, including a variety of short-term tests for genotoxicity and other relevant properties, pharmacokinetic and metabolic studies, and structure-activity relationships. These data are evaluated in combination to characterize the extent to which they support the hypothesis that an agent or chemical causes cancer in humans. The approach outlined in EPA's Guidelines for Carcinogen Risk Assessment (EPA 2005a) considers available scientific information regarding carcinogenicity and provides a narrative approach to characterizing carcinogenicity rather than assigning chemicals to specific categories (as was done previously by EPA according to the 1986 guidelines). To provide some measure of clarity and consistency in an otherwise free-form, narrative characterization, standard descriptors are used as part of the hazard narrative to express the conclusion regarding the WOE for carcinogenic-hazard potential. The five recommended standard hazard descriptors are described below.

EPA's Weight of Evidence (WOE) Descriptors (EPA 2005a)

- Carcinogenic to humans
- Likely to be carcinogenic to humans
- Suggestive evidence of carcinogenic potential
- Inadequate information to assess carcinogenic potential
- Not likely to be carcinogenic to humans

Carcinogenic to Humans: This descriptor indicates strong evidence of human carcinogenicity. This descriptor is appropriate when the epidemiologic evidence of a causal association between human exposure and cancer is convincing. Alternatively, this descriptor might be equally appropriate with a lesser weight of epidemiologic evidence that is strengthened by other lines of evidence. It can be used when all the following conditions are met: (1) evidence of an association between human exposure and either cancer or the key precursor events of the agent's mode of action is strong but insufficient for a causal association; (2) evidence of carcinogenicity in animals is extensive; (3) the mode(s) of carcinogenic action and associated key precursor events have been identified in animals; and (4) evidence is strong that the key precursor events that precede the cancer response in animals are anticipated to occur in humans and progress to tumors, based on available biological information.

Likely to Be Carcinogenic to Humans: This descriptor is appropriate when the WOE is adequate to demonstrate carcinogenic potential to humans but does not reach the WOE for the descriptor "Carcinogenic to Humans." Adequate evidence consistent with this descriptor covers a broad spectrum. At one end of the spectrum is evidence for an association between human exposure to the agent and cancer and strong experimental evidence of carcinogenicity in animals. At the other end, with no human data, the weight of experimental evidence shows animal carcinogenicity by a mode or modes of action that are relevant or assumed to be relevant to humans. The use of the term "likely" as a WOE descriptor does not correspond to a quantifiable probability. Moreover, additional data, such as information on the mode of action, might change the choice of descriptor for the illustrated examples.

Suggestive Evidence of Carcinogenic Potential: This descriptor is appropriate when the WOE is suggestive of carcinogenicity; that is, a concern for potential carcinogenic effects in humans is raised, but the data are judged not sufficient for a stronger conclusion. This descriptor covers a spectrum of evidence associated with varying levels of concern for carcinogenicity, ranging from a positive cancer result in the only study on an agent to a single positive cancer result in an extensive database that includes negative studies in other species. Depending on the extent of the database, additional studies might or might not provide further insights.

Inadequate Information to Assess Carcinogenic Potential: This descriptor is appropriate when available data are judged inadequate for applying one of the other descriptors. Additional studies generally would be expected to provide further insights.

Not Likely to Be Carcinogenic to Humans: This descriptor is appropriate when the available data are considered robust for deciding no basis for human hazard concern exists. In some instances, positive results in experimental animals can occur when the evidence is strong and consistent that each mode of action in experimental animals does not operate in humans. In other cases, there can be convincing evidence in both humans and animals that the agent is not carcinogenic. A descriptor of "not likely" applies only to the circumstances supported by the data. For example, an agent might be "Not Likely to Be Carcinogenic" by one route but not necessarily by another. In those cases that have positive animal experiment(s) but the results are judged to be not relevant to humans, the narrative discusses why the results are not relevant.

Important to note is that these WOE categories express only a **relative level of certainty that these substances might cause cancer in humans**. The categories do not specifically connote relative levels of hazard or the degree of conservatism applied in developing a dose-response assessment. For example, a substance with suggestive evidence of carcinogenic potential might impart a greater cancer risk to more people than another substance that is carcinogenic to humans.

The process of developing UREs includes several important sources of uncertainty. Many of the air toxics in NATA are classified as "likely" carcinogens. The term likely, as used in this instance, means that data are not sufficient to prove these substances definitively cause cancer in humans. That some are not human carcinogens at environmentally relevant ECs is possible, and the true cancer risk associated with these air toxics might be zero. UREs for most of the air toxics were developed from animal data using health-protective methods to extrapolate to humans. Actual human responses might differ from those predicted. For more information, see EPA's Guidelines for Carcinogen Risk Assessment (EPA 2005a).

5.2.2 Noncancer Chronic RfC

The RfC is an estimate of a **continuous inhalation exposure that is thought to be without an appreciable risk of deleterious health effects over a lifetime**. The population considered in the derivation of RfCs includes sensitive subgroups (i.e., children, asthmatics, and the elderly). The RfC is derived from the review of a health-effects database for a chemical, and identification of the most sensitive and relevant endpoint, along with the principal study or studies demonstrating that endpoint. The value is calculated by dividing the no-observed-adverse-effect level (or an analogous exposure level obtained with an alternate approach, e.g., a lowest-observed-adverse-effect level or a benchmark dose) by uncertainty factors reflecting the limitations of the data used.

The **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.

As with UREs for cancer risk assessment, the process of developing RfCs includes several important **sources of uncertainty that span perhaps an order of magnitude**. Uncertainty factors are intended to account for (1) variation in sensitivity among the individuals in the population, (2) uncertainty in extrapolating laboratory animal data to humans, (3) uncertainty in extrapolating from data obtained in a study involving a less-than-lifetime exposure, (4) uncertainty in using lowest-observed-adverse-effect-level or other data rather than no-observed-adverse-effect-level data, and (5) inability of any single study to address all possible adverse outcomes in humans adequately. Additionally, an adjustment factor is sometimes applied to account for scientific uncertainties in the data or study design not explicitly captured in the uncertainty factors (e.g., a statistically inadequate sample size or poor exposure characterization). For more information, refer to EPA's Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry (EPA 1994).

Unlike linear dose-response assessments for cancer, noncancer risks generally are not expressed as a probability that an individual will experience an adverse effect. Instead, in an air toxics risk assessment, the potential for noncancer effects in humans typically is quantified by calculating the ratio of the inhalation EC to the RfC. This ratio is referred to as the hazard quotient (HQ). For a given air toxic, exposures at or below the RfC (i.e., HQs are 1 or less) are not likely to be associated with adverse health effects. As exposures increase above the RfC (i.e., HQs are greater than 1), the potential for adverse effects also increases. The HQ, however, should not be interpreted as a probability of adverse effects. Additional information is provided in the description of risk characterization for NATA in Section 6 of this document.

5.3 Data Sources for Toxicity Values

Information on dose-response assessments for evaluating chronic exposures for NATA was obtained from multiple sources and prioritized according to OAQPS risk assessment guidelines and level of peer review. We have an approach for selecting appropriate toxicity values and, in general, this approach places greater weight on the EPA-derived toxicity values than those from other agencies (listed below). Additionally, the approach of favoring EPA values (when they exist) has been endorsed by EPA's Science Advisory Board, and it ensures the use of values most consistent with well-established and scientifically based EPA science policy. Appendix H to this document lists the toxicity values, the source of those values, and supporting information for both cancer and noncancer chronic effects used in the 2011 NATA. Cancer effects are characterized according to the extent to which available data support the hypothesis that a pollutant causes cancer in humans. Additional information on individual air toxics is included in the footnotes to the table in Appendix H.

5.3.1 U.S. EPA Integrated Risk Information System

We disseminate dose-response assessment information in several forms, depending on the level of internal review. The Integrated Risk Information System (**IRIS**) is an electronic database prepared and maintained by EPA that contains information on human-health effects that could result from exposure to various substances in the environment. These assessments have undergone external peer review and subsequent revision, compliant with requirements EPA instituted in 1996 for the IRIS review process.

Externally peer-reviewed assessments under development for IRIS were given first consideration for NATA. These assessments, which reflect the most recent available toxicity information and data analysis, were used in some cases to supersede existing values on IRIS. Current IRIS values were used for NATA when peer-reviewed IRIS values under development were not available.

5.3.2 U.S. Department of Health and Human Services, Agency for Toxic Substances and Disease Registry

The Agency for Toxic Substances and Disease Registry (**ATSDR**) publishes **minimal risk levels (MRLs)** for many substances based on **health effects other than cancer**. The MRL is defined as an estimate of human exposure to a substance that is likely to be without an appreciable risk of adverse effects (other than cancer) over a specified duration of exposure. For noncancer values in the 2011 NATA, inhalation MRLs were used when IRIS RfC values were not available or when the ATSDR value was based on more recent, peer-reviewed data and analysis methods than the IRIS value, because the ATSDR concept, definition, and derivation are analogous to IRIS. ATSDR does not develop assessments based on carcinogenicity. After internal and external review, MRLs are published in pollutant-specific toxicological-profile documents. ATSDR regularly updates these toxicological-profile documents and they are available at Toxic Substances Portal MRLs (ATSDR 2015).

5.3.3 California Environmental Protection Agency Office of Environmental Health Hazard Assessment

California's Office of Environmental Health Hazard Assessment (**OEHHA**) develops **UREs** based on carcinogenicity and reference exposure levels (**RELs**) based on health effects other than cancer. The REL is defined as a concentration level at or below which no adverse health effects are anticipated. For cancer and noncancer values in the 2011 NATA, OEHHA UREs and inhalation RELs were used when their derivation was determined to be consistent with the concepts and definitions of IRIS or ATSDR. OEHHA dose-response information is available at Air Toxicology and Epidemiology (OEHHA 2014). Technical support documents for assessing hot spots are available on the OEHHA website at Hot Spots Guidelines (OEHHA 2015).

5.3.4 U.S. EPA Health Effects Assessment Summary Tables

The Health Effects Assessment Summary Tables (EPA 2008a) are a comprehensive listing consisting almost entirely of **provisional** UREs, RfCs, and other risk assessment information of interest that various EPA offices have developed. The assessments, which have never been submitted for EPA consensus, were last updated in 2001. NATA uses information from these tables only when no values from the sources discussed in Sections 5.3.1 through 5.3.3 are available.

5.3.5 World Health Organization International Agency for Research on Cancer

The International Agency for Research on Cancer of the World Health Organization (WHO) coordinates and conducts research on cancer and provides information on related cancer research and epidemiology. Although the agency does not develop quantitative dose-response values, it has published a series of monographs (WHO 2015) on the carcinogenicity of a wide range of substances. The following "degrees of evidence" the International Agency for Research on Cancer has published are included in Appendix H to this document as supporting information when EPA WOE determinations were not available for a substance or are out of date:

Group 1:	Carcinogenic to humans;
Group 2A:	Probably carcinogenic to humans;
Group 2B:	Possibly carcinogenic to humans;
Group 3:	Not classifiable as to human carcinogenicity; and
Group 4:	Probably not carcinogenic to humans.

5.4 Additional Toxicity Decisions for Some Chemicals

After the dose-response information was prioritized, we made additional changes to some of the chronic inhalation exposure values to address data gaps, increase accuracy, and avoid underestimating risk for NATA. Important changes made for the 2011 NATA are outlined below and are reflected in Appendix H to this document.

5.4.1 Polycyclic Organic Matter

A substantial proportion of polycyclic organic matter (POM) reported in the 2011 NEI was not speciated into individual compounds. For example, some emissions of POM were reported in NEI as "7-PAH" or "16-PAH," representing subsets of certain POM, or simply as "total PAH" or "polycyclic organic matter." In other cases, individual POM compounds were reported for which no quantitative cancer dose-response value has been published in the sources used for NATA. As a result, simplifying assumptions that characterize emissions reported as POM were applied so that cancer risk could be quantitatively evaluated for these species without substantially under- or overestimating risk (which can occur if all reported emissions of POM were assigned the same URE). To accomplish this, **POM emissions as reported in NEI were grouped into categories**. EPA assigns dose-response values based on the known or estimated toxicity for POM within each group and on information for the POM speciation of emission sources, such as wood fires and industrial processes involving combustion.

For the 2011 NATA, unspeciated POM emissions were divided into **eight POM groups**. The first two groups included unspeciated POM (including "total PAH") and individual POM species with no URE assigned. Both groups were assigned a URE equal to 5 percent of that for pure benzo[a]pyrene taking into account toxicity and the estimated emission profile of POM compounds. Groups 3 through 7 comprised POM compounds, the emissions for which were reported as individual compounds and for which UREs have been estimated. Compounds in these groups were categorized based on toxicity, and an appropriate URE was assigned to each category based on toxicity of the compounds included in the group. Category 8 was composed of unspeciated polynuclear aromatic hydrocarbons reported as 7-PAH and was assigned a URE equal to 18 percent of that for pure benzo[a]pyrene. We discuss the POM groups used for the 2011 NEI in Section 2.1.1.2.

We concluded that three PAHs—anthracene, phenanthrene and pyrene—are not carcinogenic and therefore no URE was assigned for the 2011 NATA. Details of the analysis that led to this conclusion can be found in the document entitled Development of a Relative Potency Factor (RPF) Approach for Polycyclic Aromatic Hydrocarbon (PAH) Mixtures: In Support of Summary Information of the Integrated Risk Information System (IRIS) (EPA 2010a).

5.4.2 Glycol Ethers

Much of the emission-inventory information for the glycol ether category reported only the total mass for the entire group without distinguishing among individual glycol ether compounds. In other cases, emissions of individual glycol ether compounds that had not been assigned dose-response values were reported. Individual glycol ether compounds vary substantially in toxicity. To avoid underestimating the health hazard associated with glycol ethers, we protectively applied the RfC for **ethylene glycol methyl ether** (the most toxic glycol ether for which an assessment exists) to glycol ether emissions of unspecified composition.

5.4.3 Metals

Several decisions made for the 2011 NATA regarding the toxicity values used for metal compounds are discussed in this section.

Chromium (VI) compounds. The IRIS RfC for particulate chromium (VI) was used instead of the RfC for chromic acid mists and dissolved aerosols to avoid underestimating the health hazard associated with these compounds. The RfC for particulate chromium (VI) is less than those RfCs for chromic acid mists and dissolved aerosols.

Lead. We consider the primary National Ambient Air Quality Standard (NAAQS) for lead, which incorporates an ample margin of safety, to be protective of all potential health effects for the most susceptible populations. The NAAQS, developed using the EPA Integrated Exposure, Uptake, Biokinetic Model, was preferred over the RfC for noncancer adverse effects because the NAAQS for lead was developed using more recent toxicity and dose-response information on the noncancer adverse impacts of lead. The NAAQS for lead was set to protect the health of the most susceptible children and other potentially at-risk populations against an array of adverse health effects, most notably including neurological effects, particularly neurobehavioral and neurocognitive effects (which are the effects to which children are most sensitive). The lead NAAQS, a rolling 3-month average level of lead in total suspended particles, was used as a long-term value in NATA.

Nickel compounds. The cancer inhalation URE for most of the emissions of nickel compounds included in NATA (including unspecified nickel emissions reported as "nickel compounds") was derived from the IRIS URE for insoluble nickel compounds in crystalline form. Soluble nickel species, and insoluble species in amorphous form, do not appear to produce genotoxic effects by the same toxic mode of action as insoluble crystalline nickel. Nickel speciation information for some of the largest nickel-emitting sources, including oil and coal combustion, suggests that at least 35 percent of total nickel emissions could be soluble compounds. The remaining insoluble nickel emissions, however, are not well characterized. Consistent with this limited information, we conservatively assumed for NATA that 65 percent of emitted nickel is insoluble and that all insoluble nickel is crystalline. Because the nickel URE listed in IRIS is based on nickel subsulfide and represents pure insoluble crystalline nickel, it was adjusted to reflect an assumption that 65 percent of the total mass of emitted nickel might be carcinogenic. In cases where a chemical-specific URE was identified for a reported nickel compound, it was used without adjustment. Furthermore, the MRL in Table 2 of the ATSDR is not adjusted because the noncancer effects of nickel are not thought to be limited to the crystalline, insoluble form.

5.4.4 Adjustment of Mutagen UREs to Account for Exposure During Childhood

For carcinogenic chemicals acting via a mutagenic mode of action (i.e., chemicals that cause cancer by damaging genes), we recommend that estimated risks reflect the increased carcinogenicity of such chemicals during childhood. This approach is explained in detail in the Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens (EPA 2005c). Where available data do not support a chemical-specific evaluation of differences between adults and children, the Supplemental Guidance recommends using the following default **adjustment factors for early-life exposures**: increase the carcinogenic potency by 10-fold for children up to 2 years old and by 3-fold for children 2 to 15 years old. These adjustments have the aggregate effects of increasing by about 60 percent the estimated risk (a 1.6-fold increase) for a lifetime of constant inhalation exposure. EPA recommends that these default adjustments be made only for carcinogens known to be mutagenic for which data to evaluate adult and juvenile differences in toxicity are not available.

For NATA 2011, the UREs for acrylamide, benzidine, chloroprene, coke oven emissions, ethyl carbamate, methylene chloride, nitrosodimethylamine, and PAHs were adjusted upward, by multiplying by a factor of 1.6, to account for the increased risk during childhood exposures. Although trichloroethylene is carcinogenic by a mutagenic mode of action, the age-dependent adjustment factor for the URE only applies to the portion of the slope factor reflecting risk of kidney cancer. For full lifetime exposure to a constant level of trichloroethylene exposure, the URE was adjusted upward by a factor of 1.12 (rather than 1.6 as discussed above). For more information on applying age-dependent adjustment factors in cases where exposure varies over the lifetime, see Toxicological Review of Trichloroethylene (EPA 2014c). These air toxics are the only ones that met the criteria described in the previous paragraph at the time of this assessment. The overall lifetime adjustment was applied because a single, lifetime-average EC was estimated for NATA rather than age-group-specific exposures. The URE for vinyl chloride includes exposure from birth, although the IRIS assessment contains UREs for both exposure from birth and exposure during adulthood. This value already accounts for childhood exposure; thus, no additional factor was applied.

5.4.5 Diesel Particulate Matter

EPA uses an IRIS RfC for adverse noncancer effects of diesel PM. Recently, several large epidemiology studies (Attfield, et al. 2012; Garshick, et al. 2012; Silverman, et al. 2012) have been published that strengthen the WOE that diesel exhaust is carcinogenic to humans. Two of these studies included quantitative estimates of exposure. Partly on the basis of these studies, the International Agency for Research on Cancer elevated its classification of diesel exhaust to **"carcinogenic to humans"** (Group 1) in 2012 (IARC 2013). We requested the Health Effects Institute (HEI) evaluate the suitability of the new epidemiology studies for developing a cancer potency. In November 2015, HEI published its report on these new studies (HEI 2015) and concluded that they are sufficiently robust to estimate quantitative cancer risks and estimate uncertainties. At this time, a URE for diesel PM, based on these current studies, has not yet been derived and a quantitative assessment of the cancer risk has not been included in the 2011 NATA. Evidence exists, however, that the general population is exposed to levels close to or overlapping with apparent levels that have been linked to increased cancer risk in epidemiological studies. Based on the Health Assessment Document for Diesel Engine Exhaust (EPA 2002a), we concluded that national-average lifetime cancer risk from exposure to diesel exhaust (which contains both gases and particulate matter) could exceed 1 in 100,000 and could be as high as 1 in 1,000, although the lower end of the risk range includes zero.

5.5 Summary

- To evaluate the potential of a given air toxic to cause cancer and other adverse health effects, we identified potential adverse effects that a particular substance causes and evaluated the specific ECs at which these effects might occur.
- The URE represents the upper-bound excess cancer risk estimated to result from continuous exposure to a concentration of 1 µg of a substance per m^3 of air over a 70-year lifetime.
- The RfC is an estimate of a continuous inhalation EC over a 70-year lifetime that is thought to be without an appreciable risk of deleterious effects. The population considered in the derivation of RfCs includes sensitive subgroups (i.e., children, asthmatics, and the elderly).
- Dose-response-assessment information for chronic exposure was obtained from multiple sources and prioritized according to conceptual consistency with OAQPS risk assessment guidelines and level of peer review.

- After considering dose-response information, EPA adjusts some chronic-toxicity values to increase accuracy and to avoid underestimating risk.

6 CHARACTERIZING RISKS AND HAZARDS IN NATA

Risk characterization, the final step in our risk assessment process for air toxics, combines the information from modeled exposure estimates with the dose-response assessment, providing a **quantitative estimate of potential cancer risk and noncancer hazard** associated with real-world exposure to air toxics. The term "risk" implies a statistical probability of developing cancer over a lifetime. Noncancer "risks," however, are not expressed as a statistical probability of developing a disease. Rather, noncancer "hazards" are expressed as a ratio of the EC to an RfC associated with observable adverse health effects (i.e., an HQ).

This section contains information on the risk characterization conducted for NATA. After a brief overview of the risk-related questions that NATA is intended to address, the methods used to conduct characterization of cancer risk and noncancer hazards for NATA are described. A discussion of the quantitative results included in NATA follows this description.

6.1 The Risk-characterization Questions NATA Addresses

The NATA risk characterization considers both cancer risk and the potential for noncancer effects from inhalation of air toxics nationwide, in both urban and rural areas. The purpose of NATA is to **understand cancer risks and noncancer hazards** to help EPA and others identify air toxics and source categories of greatest potential concern and to set priorities for collecting additional information to improve future assessments. The assessment represents **a "snapshot" in time for characterizing risks from exposure to air pollutants**; it is not designed to characterize risks sufficiently for regulatory action. The risk characterization for NATA, which was limited to inhalation risk from outdoor sources, was designed to answer the following questions:

- Which air toxics pose the greatest potential risk of cancer or adverse noncancer effects across the entire United States?
- Which air toxics pose the greatest potential risk of cancer or adverse noncancer effects in specific areas of the United States?
- Which air toxics pose less, but still significant, potential risk of cancer or adverse noncancer effects across the entire United States?
- When risks from inhalation exposures to all outdoor air toxics are considered in combination, how many people could experience a lifetime cancer risk greater than levels of concern (e.g., 1-in-1 million)?
- When potential adverse noncancer effects from long-term exposures to all outdoor air toxics are considered in combination for a given target organ or system, how many people could experience exposures that exceed the reference levels intended to protect against those effects (i.e., a hazard quotient greater than 1)?

6.2 How Cancer Risk is Estimated

To estimate cancer risks in NATA, the results of cancer dose-response assessments for a given chemical were converted to a URE that is then multiplied by the estimated inhalation EC to obtain an estimate of individual lifetime cancer risk. The approach used in NATA for characterizing cancer risk is consistent with EPA's 2005 final Guidelines for Carcinogen Risk Assessment (EPA 2005a). When used in conjunction with the cancer UREs described in Section 5, the approach is also consistent with EPA's associated documentation on Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens (EPA 2005c).

6.2.1 Individual Pollutant Risk

Individual lifetime cancer risk associated with exposure to a single air pollutant was estimated by multiplying an average estimated long-term EC by the corresponding URE for that pollutant. Thus, the below equation estimates the probability of an individual developing cancer over a lifetime due to a given inhalation exposure.

$$Risk = EC \times URE$$

Where:

Risk = estimated incremental lifetime cancer risk for an individual as a result of exposure to a specific air toxic, unitless (expressed as a probability)

EC = estimate of long-term inhalation exposure concentration for a specific air toxic, in units of $\mu g/m^3$

URE = the corresponding inhalation unit risk estimate for that air toxic, in units of $1/(\mu g/m^3)$

Of importance to note is that UREs are typically upper-bound estimates, so actual risks might be lower than predicted. Also, the true value of the risk is unknown.

6.2.2 Multiple-pollutant Risk

The individual lifetime cancer risk resulting from exposure to multiple air toxics was estimated by summing the chronic cancer risk for each air toxic that can be quantified. This estimate of risk focused on the additional lifetime risk of cancer predicted from the exposure being analyzed, over and above that due to any other factors. The following equation estimates the predicted cumulative individual cancer risk from inhalation of multiple substances:

$$Risk_{tot} = Risk_1 + Risk_2 + \cdots + Risk_i$$

Where:

$Risk_{tot}$ = total cumulative individual lifetime cancer risk, across *i* substances

$Risk_i$ = individual risk estimate for the *i*th substance

For NATA, the estimated ECs were not considered upper bounds. Rather, they represented central-tendency estimates of ECs for each demographic group at the geographic unit of analysis (e.g., the census-tract level). Because cancer slope factors were 95-percent upper-confidence intervals (not "most probable estimates"), summing traditional risk levels can cause the resulting sum to overestimate a 95-percent upper-confidence-level risk for a mixture.

The NATA approach assumed an additive effect from simultaneous exposures to several carcinogens. **Summing cancer risk estimates** is not appropriate when effects from multiple chemicals are synergistic (greater than additive) or antagonistic (less than additive). Notwithstanding the statistical limitations of summing traditional risk estimates and the implicit assumption that the toxicities will be additive (i.e., no interactions such as synergism or antagonism occur), the numerical ease for combining risk in this way makes this method the most popular for approximating cumulative risks in the short term, at least for a screening level of assessment. Information on non-additive interactions is not readily available in a form that can be used for NATA. In the absence of specific information, therefore, cancer risk from various chemicals is conservatively assumed to be additive. Thus, the cancer risks from all air toxic compounds listed as carcinogenic or likely carcinogenic to humans were summed to determine cumulative cancer risks for NATA. More information on EPA's methods for conducting risk assessment of mixtures can be found in the Framework for Cumulative Risk Assessment (EPA 2003).

6.3 How Noncancer Hazard is Estimated

To evaluate the potential for noncancer adverse health effects, chronic dose-response data were used to estimate a threshold that is the EC at which adverse health effects are assumed to be unlikely (i.e., the RfC). (See Section 5.2.2 for more information on noncancer RfCs.) Due to the wide variety of endpoints, hazard-identification procedures for noncancer effects have not been described as completely in EPA guidance as procedures for the identification of carcinogens. EPA has published guidelines, however, for assessing several specific types of chronic noncancer effects (mutagenicity, developmental toxicity, neurotoxicity, and reproductive toxicity) that can be found at Products and Publications Relating to Risk Assessment Produced by the Office of the Science Advisor (EPA 2015l). EPA has also published a framework for using studies of these and other effects in inhalation risk assessment (EPA 1994).

EPA's Chronic Noncancer Guidelines

- Mutagenicity (EPA 1986)
- Developmental Toxicity (EPA 1991)
- Neurotoxicity (EPA 1998)
- Reproductive Toxicity (EPA 1996)

6.3.1 Individual Pollutant Hazard

Chronic noncancer hazards were estimated for NATA by dividing a chemical's estimated long-term EC by the RfC for that chemical to yield an HQ. The following equation estimates the noncancer hazard due to a given inhalation exposure:

$$HQ = \frac{EC}{RfC}$$

Where:

HQ = the hazard quotient for an individual air toxic, unitless

EC = estimate of long-term inhalation exposure concentration for a specific air toxic, in units of mg/m^3

RfC = the corresponding reference concentration for that air toxic, in units of mg/m^3

An HQ value less than or equal to 1.0 indicates that the exposure is not likely to result in adverse noncancer effects. An HQ value greater than 1.0, however, does not necessarily suggest a likelihood of adverse health effects and cannot be interpreted to mean that adverse health effects are statistically likely to occur. The statement is simply whether, and by how much, an EC exceeds the RfC, indicating that a potential exists for adverse health effects.

6.3.2 Multiple-pollutant Hazard

Chronic noncancer hazards for multiple air toxics were estimated by summing chronic noncancer HQs for individual air toxics that cause similar adverse health effects to yield a hazard index (HI). Aggregation in this way produces a target-organ-specific HI, defined as a sum of HQs for individual air toxics that affect the same organ or organ system. More information on chemical mixtures risk assessment methods can be found in the EPA supplementary guidance for risk assessment of mixtures (EPA 2000).

The following equation estimates the HI from inhalation of multiple substances:

$$HI = HQ_1 + HQ_2 + \cdots + HQ_i$$

Where:

HI = the hazard index for chronic exposure to air toxics 1 through *i*, unitless

HQi = the hazard quotient for the *i*th air toxic, where all *i* air toxics are assumed to affect the same target organ or organ system, unitless

As with the HQ, an HI value less than or equal to 1.0 indicates that the exposure is not likely to result in adverse noncancer effects. An HI value greater than 1.0, however, does not necessarily suggest a likelihood of adverse health effects and cannot be interpreted as a statistical probability of adverse effects occurring.

This equation assumes an additive effect from simultaneous exposures to several chemicals. **Summing of HQs** is inappropriate when effects from multiple chemicals are synergistic (greater than additive) or antagonistic (less than additive). As is the case with cancer risk, quantitative information on non-additive interactions resulting in noncancer hazards is not readily available; consequently, the noncancer HQs are assumed to be additive for chemicals with the same target organ or organ system. For NATA, noncancer hazards could be combined for six target organs or systems: respiratory, cardiovascular, blood, liver/kidney, nervous, and immune. Results from the assessment indicated that the primary noncancer hazards for inhalation exposures to the modeled chemicals were respiratory hazards. As a result, the 2011 assessment presents noncancer results for all target organ endpoints in the form of HQs; **HIs are reported only for respiratory endpoints**.

6.4 How Risk Estimates and Hazard Quotients are Calculated for NATA at Tract, County, and State Levels

The cancer risk and HQs for each modeled air toxic are estimated from ECs (not ambient concentrations) by combining them with UREs and inhalation RfCs (or their equivalents). As described previously, the modeling conducted for NATA resulted in ambient concentrations for each air toxic emitted by modeled sources, with the level of spatial resolution varying by source type and the corresponding modeling approach (see Section 3). NATA point, nonpoint, mobile onroad, and mobile nonroad sources were modeled at the census-block level in HEM-3. However, nonpoint, mobile onroad, and mobile nonroad (except CMV) emissions were allocated from county to the census-tract level. Secondary formation, fires (wild, prescribed and agricultural field burning) and primary biogenic concentrations were estimated at the 12km grid level using CMAQ.

6.4.1 Model Results for Point Sources: Aggregation to Tract-level Results

HEM-3 was used to estimate ambient concentrations for point-source emissions, and model results were generated at the block level. For risk and exposure calculations, we aggregated concentration results to the tract level by taking a **population-weighted average** of all of the block-level concentrations within a given tract, as follows:

$$Conc_{tract\ i} = \frac{\sum Pop_{block\ j} \times Conc_{block\ j}}{\sum Pop_{block\ j}}$$

Where:

$Conc_{tracti}$ = ambient concentration for census tract *i*

$Conc_{block\ j}$ = ambient concentration for census block *j* (contained within tract *i*), estimated by HEM-3

$Pop_{block\ j}$ = population of blocks contained in tract *i*

Unweighted average concentrations also were calculated at the tract level as follows:

$$Conc_{tract\ i} = \frac{\sum Conc_{block\ j}}{n}$$

Where:

$Conc_{tracti}$ = ambient concentration for census tract *i*

$Conc_{blockj}$ = ambient concentration for census block *j* (contained within tract *i*), estimated by HEM-3

n = number of census blocks contained in tract *i*

6.4.2 Background Concentrations and Secondary Pollutants: Interpolation to Tract-level Results

Background concentrations, as well as estimated concentrations of secondary pollutants generated by the CMAQ model, were estimated for levels other than census tract and thus required interpolation "down" to the tract level. Background concentrations were estimated at the county level. To obtain tract-level concentrations, the county-level estimate was assigned to all census tracts within that county. For secondary pollutants, concentrations were estimated using CMAQ. The results for each grid were then **applied evenly** to all tracts located within the grid.

6.4.3 Aggregation of Tract-level Results to Larger Spatial Units

Tract-level ambient concentrations were aggregated up to the county, state, regional, and national level using a method that weights concentration according to the population within a region. For a county, for example, a **population-weighted ambient concentration** was estimated by multiplying the tract-level concentrations by the population of each tract, summing these population-weighted concentrations, and dividing by the total county population encompassing all tracts to obtain a final population-weighted, county-level concentration. The process for aggregating from the tract to the county level can be expressed using the following equation:

$$Conc_{county\ k} = \frac{\sum(Conc_{tract\ i} \times Pop_{tract\ i})}{Pop_{county\ k}}$$

Where:

$Conc_{countyk}$ = population-weighted concentration for county *k*

$Conc_{tracti}$ = ambient concentration in tract *i* (contained within county *k*)

Pop_{tracti} = population in tract *i* (contained within county *k*)

$Pop_{countyk}$ = population in county *k*

This same method was applied when aggregating up to the state, regional, or national level, using the appropriate concentration and population values. NATA includes ambient concentrations, ECs, cancer risks, and noncancer HQs at the tract, county, state, regional, and national levels.

The ambient concentrations derived at the block level also were used to estimate ECs using **either direct exposure modeling with HAPEM or with the exposure factors derived from the HAPEM modeling** (i.e., ratios of EC to estimated ambient concentration). (See Section 4 for a more thorough discussion of NATA exposure modeling and estimates.) Because the exposure factors were applied at the tract level, each census block was assigned the tract-level EC or exposure factor and then the census-block-level ECs are estimated. As was done with the ambient-level concentrations, the block-level ECs were used to estimate cancer and noncancer effects and to aggregate these concentrations up to larger spatial scales. To aggregate tract-level concentrations up to the county-, state-, regional-, or national-level concentrations, the tract-level concentrations were population-weighted.

6.5 The Risk Characterization Results that NATA Reports

NATA provides a snapshot of the outdoor air quality and the risks to human health that would result if air toxic emission levels remain unchanged. The assessment was based on an inventory of air toxics emissions from 2011. **Individuals were assumed to spend their entire lifetimes exposed to these air toxics**. Therefore, the reductions in emissions that have occurred since the year of the assessment, or those that might happen in the future due to regulations for mobile and industrial sources, were not accounted for. Each NATA represents an update and enhancement to the previous NATA. Because improvements in methodology are made with each successive assessment, **comparing assessment results from year to year is not meaningful**. Any change in emissions, ambient concentrations, or risks might be due to either improvement of methodology or to real changes in emissions or source characterization.

The evaluation of national-scale results and comparison of risks among chemicals make it possible to estimate which air toxics pose the greatest potential risk to human health in the United States. NATA reports a summary of these findings. Cancer risks are presented as lifetime risks, meaning the risk of developing cancer as a result of inhalation exposure to each air toxic compound over a normal lifetime of 70 years. Noncancer hazards are presented in terms of the ratio between the exposure and an RfC for inhalation exposures (i.e., the HQ). As described previously in this section, HQs are combined across chemicals where a common target organ or system is expected to estimate HI (i.e., for respiratory).

Using these quantitative results, **NATA classifies certain pollutants as drivers or contributors** at the national or regional scale based on certain criteria. Exhibit 64 contains the criteria for classifying the air toxics included in NATA at the regional and national level. In general, drivers and contributors were defined as air toxics showing a particular level of risk or hazard for some number of people exposed.

Exhibit 64. NATA Drivers and Contributors of Health Effects for Risk Characterization

	Criterion (Criteria in both columns must be met)	
Risk-characterization Category	**Individual Health Risk or Hazard Index Exceeds…**	**Minimum Number of People Exposed (in millions) is…**
Cancer Risk *(value in first column represents individual lifetime cancer risk, in 1 million)* [a]		
National cancer driver	10	25
Regional cancer driver (either set of criteria can be used)	10	1
	100	0.01
National cancer contributor	1	25
Regional cancer contributor	1	1
Hazard Index *(value in first column represents chronic hazard index for any organ/organ system)* [b]		
National noncancer driver	1.0	25
Regional noncancer driver	1.0	0.01

[a] Cancer risks are upper-bound lifetime cancer risks; that is, a plausible upper limit to the true probability that an individual will contract cancer over a 70-year lifetime as a result of a given hazard (such as exposure to a toxic chemical). This risk can be measured or estimated in numerical terms (e.g., one chance in a hundred).

[b] Hazard index is the sum of the HQs for substances that affect the same target organ or organ system. Because different pollutants can cause similar adverse health effects, combining HQs associated with different substances is often appropriate to understand the potential health risks associated with aggregate exposures to multiple pollutants.

For example, for a pollutant to be categorized in NATA as a cancer contributor at the national level, the individual lifetime cancer risk for that pollutant must have been shown by the assessment to be 1-in-1 million *and* the number of people exposed to that pollutant must have been shown to be at least 25 million. For a pollutant to be categorized in NATA as a regional driver of noncancer health effects, the chronic hazard index for that pollutant must have been shown to exceed 1.0 *and* the number of people exposed to that pollutant must have been shown to be at least 0.01 million.

The NATA results for 2011 indicated that **most individuals' estimated risk was between 1-in-1 million and 100-in-1 million**, although a small number of localized areas showed risks of higher than 100-in-1 million. Although individuals and communities might be concerned about these results, recall that **NATA was not designed to assess specific risk values at local levels**. The results are best used as a tool to prioritize pollutants, emissions sources, and locations of interest for further investigation. Furthermore, readers are reminded that the risks estimated by the assessment do not consider indoor sources of air toxics or ingestion exposure to any pollutants. Also, although NATA estimates cancer and noncancer risks for numerous pollutants, additional chemicals might exist that are not identified or for which toxicity information is unavailable. Therefore, these risk estimates represent only a subset of the total potential cancer and noncancer risk associated with air toxics.

Analytical results (including modeled ambient concentrations, exposure, and risks) for each NATA are also provided at the census-tract, county, and state level for those who wish to do their own technical analyses using the most refined output available. **The results from all NATAs can be found here.** In performing such analyses, users must be extremely mindful of the purposes for which NATA was developed. NATA was developed as a tool to inform both national and more localized efforts to collect air toxics information and characterize emissions (e.g., prioritize pollutants or geographic areas of interest for more refined data collection such as monitoring). The results are most meaningful when viewed at the state or national level. Nevertheless, reported spatial patterns within a county likely represent actual variations in *overall average* population risks. Less likely, however, is that the assessment pinpoints the exact locations where higher risks exist or that the assessment captures the highest risks in a county. Using these results alone to draw conclusions about local concentrations and risk is inappropriate.

This assessment did not focus on the identification of geographic areas or populations that have significantly higher risks than others. Rather, it focused on characterizing geographic patterns and ranges of risk across the country. In general, however, spending time in larger urban areas tends to pose greater risks than spending time in smaller urban and rural areas because the emissions of air toxics tend to be higher and more concentrated in areas with more people. This trend is not, however, universal and can vary from pollutant to pollutant according to its sources. The trend also can be affected by exposures and risk from non-inhalation and indoor sources of exposure.

Based on the NATA results, millions of people live in areas where air toxics pose potential health concerns. Although air quality continues to improve, **more needs to be done to meet the CAA's requirements to reduce the potential exposure and risk from these chemicals**. We will continue to develop air toxic regulations and cost-effective pollution prevention and other control options to address indoor and urban pollutant sources that significantly contribute to risk.

6.6 Summary

- The purpose of NATA is to understand cancer risk and noncancer health effects to help EPA and others identify pollutants and source categories of greatest potential concern and to set priorities for collecting additional information to improve future assessments.
- Cancer risk was expressed as a statistical probability that an individual will develop cancer. Cancer risks were assumed to be additive across chemicals for NATA.
- Noncancer hazard was expressed as an HQ, which is the ratio of the EC to an RfC associated with observable adverse effects.
- NATA estimated most individuals' risk to be between 1-in-1 million and 100-in-1 million, although a small number of localized areas showed risk higher than 100-in-1 million.
- Air toxics data for NATA are presented at the national, regional, state, county, and census-tract levels. The results are most meaningful when viewed at the state or national level. Using these results in the absence of additional information to draw conclusions about local concentrations and risk is inappropriate.

7 VARIABILITY AND UNCERTAINTY ASSOCIATED WITH NATA

7.1 Introduction

Presented in this section are discussions on variability and uncertainty associated with the NATA process. Clearly understanding these two fundamental concepts—inherent in all broad-scale assessments that rely on models and data—will enable the users of the NATA results to understand which questions can be answered appropriately and which cannot.

Key Definitions for this Section

Variability represents the ***diversity or heterogeneity in a population or parameter*** (e.g., variation in heights of people). Variability cannot be reduced by taking more (or better) measurements; however, it can be accounted for by a more detailed modeling approach (e.g., modeling peoples' heights in terms of age will reduce the unexplained variability due to variation in heights).

Uncertainty refers to the ***lack of knowledge*** regarding the actual values of model input variables (parameter uncertainty) and of physical systems (model uncertainty). Uncertainty can be reduced through improved measurements and improved model formulation.

As stated in Section 1, NATA results should not be used for limited-scale or site-focused applications. NATA results are intended to characterize broad-scale risk to help identify those air toxics and source types associated with the highest exposures and posing the greatest potential health risks. The results are intended to identify geographic patterns and ranges of risks across the country. To avoid over-interpretation and misapplication of the results, users must first understand the concepts of variability and uncertainty and then must recognize the role that these elements play in the NATA results.

Air toxic emissions, air concentrations, and exposures are not the same throughout the United States, and the risks associated with air toxics are not the same for all people. Some geographic areas have higher concentrations than others. At certain times, the concentration is higher at a given location than at other times. The risks for some individuals are below the national average, while for others the risks are above the national average. For these reasons, understanding how the ambient (outdoor) air concentration, exposure, and risk from air toxics vary throughout the United States is essential for understanding NATA. This information comes from a process called variability analysis.

EPA seeks to protect health with reasonable confidence based on the best data available. Estimates of air concentrations, exposures, and risks, however, necessarily always involve assumptions. Assumptions are necessary to simplify the problem at hand, while also making assessment possible given available information and resources. Assumptions introduce uncertainties into the results because confidence that the assumptions are entirely correct is not possible. Understanding the magnitude of these uncertainties, the level of confidence that can be placed in statements related to the assessment, and how this confidence affects the ability to make reasoned decisions is essential. This information comes from a process called uncertainty analysis.

7.2 How NATA Addresses Variability

The NATA process focuses on the variation in ambient air concentrations, exposures, and risks in geographic areas of the United States, Puerto Rico, and the U.S. Virgin Islands. Included, for example, are variations in the locations of various sources and the amounts of pollutants that these sources emit, variations in meteorological conditions in various parts of the country, and variations in the daily activities of people. This section presents information on the key components that drive variability in risks associated with air toxics and the variability components that NATA addresses. A brief explanation is also provided on how NATA results should be interpreted in light of variability.

7.2.1 Components of Variability

The NATA results show how air concentrations, exposures, and risks vary across broad geographic regions of the country. They do not fully characterize how concentration, exposure, and risk vary among individuals, except to the extent these individuals live in different geographic regions and are affected by the values typical of a census tract in that region. NATA results also do not fully characterize how ambient air concentrations might vary temporally and they do not characterize how concentrations vary spatially within a census tract. The following list contains explanations of some of the components of variability that determine differences in ambient air concentrations and individual risks. Key components driving variability in risk associated with air toxics include temporal variation, geographic variation, and variations in where people live, their levels of activities, and their degrees of susceptibility or sensitivity, as described below.

Temporal. Sources do not emit pollutants at constant rates. Similarly, the meteorological conditions that affect dispersion in the atmosphere vary over time. Thus, the ambient air concentration at a given location can vary over time.

Geographic. The influence of pollutant emissions on ambient concentrations at a particular location depends on the degree of atmospheric dispersion of the emissions as they travel from the source to the receptor. Dispersion depends on both meteorological conditions, which vary from place to place, and the travel distance from source to receptor. As a result, the ambient air concentration can vary greatly among different locations. The NATA analysis accounts for some geographic variation by using available meteorology data representative of the location and by modeling ambient concentrations for census areas, but the spatial resolution of model predictions is limited.

Individual location. Two individuals might live at different locations within the same census tract. The ambient concentration estimated for the tract is only an approximation of conditions at all locations in the tract. Different locations within that tract might have different average ambient concentrations. Therefore, exposures and risks also can vary.

Individual activity patterns. Two individuals might live at the same location but engage in different activities (called an "activity pattern") during each day. Concentrations of substances indoors often differ from concentrations outdoors. If one person spends more time indoors than the other person does, the average air concentration to which the two are exposed will differ, even though the ambient air concentration is the same. Similarly, one person might spend more time in a car than the other person might and be exposed to an air concentration that is typical near roads. The net effect would be that the concentration of each pollutant in the air actually inhaled by these two individuals would differ. In other words, the exposure differs for these two individuals.

In addition, buildings and vehicles vary with respect to the amount of outdoor pollution that penetrates into the indoor and in-vehicle microenvironments due to differences in ventilation and building and vehicle integrity. Thus, two people who live in the same location and spend the same amount of time indoors can still be exposed to different pollutant concentrations.

Susceptibility. Two individuals might live at the same location and engage in the same activities, but one person might be more susceptible than another might be. Susceptibility refers to the extent to which an individual takes a pollutant into the body, transports it into an organ or tissue that might be adversely affected by it, or develops an adverse effect.

An individual who is more susceptible might develop a higher concentration of a pollutant in his or her organs or tissues, or have a higher chance of developing an adverse health effect, than another individual even though the exposures for both individuals are the same. For example, people breathe at different rates; two individuals placed into exactly the same air might bring different amounts of a pollutant into their bodies. The amount of a pollutant reaching an organ or tissue also might vary from individual to individual, even if both bring the same amount into their lungs. The amount of time the pollutant remains in the body also might differ. Finally, the innate sensitivity to the effect might vary even at equal doses in the tissues. The net effect of these factors is that either the dose of the pollutant delivered to the organs or tissues of the body or the level of response, or both, can differ substantially between these two individuals, even though the individuals are exposed to exactly the same pollutant concentrations.

The extent to which each factor described above influences variation in individual risk can depend on the age, gender, or ethnic group to which an individual belongs, as well as on that individual's lifestyle. These groups comprise different receptor populations, or cohorts, and the exposures and risks can differ among them.

7.2.2 Quantifying Variability

EPA conducts NATA to understand how ambient air concentration, exposure, and risk vary geographically and not among specific individuals. EPA calculates the ambient air concentrations for each specific, discrete location (i.e., census-block centroid or census-tract centroid; see discussion below) based on the emission sources and meteorological conditions affecting those specific tracts. Some temporal variation is accounted for in NATA calculations. For example, meteorology data used for air quality modeling is temporally dynamic. The air quality modeling therefore captures important variations in ambient conditions on an hourly basis before the resulting modeled ambient air concentrations are time-averaged. The ambient concentration inputs to HAPEM are stratified into eight 3-hour time blocks; HAPEM then calculates ECs for each 3-hour time block before calculating an overall, long-term average EC. Although this approach to air quality and exposure modeling takes into account some important temporal variations, these time-stratified model outputs are averaged prior to the risk characterization step and are not included in the NATA results reported by EPA.

The NATA concentrations and risks, however, do reflect a degree of geographic variation. The smallest geographic area for which NATA results are reported is the census tract. Although results are reported at the census-tract level, average risk estimates are far more uncertain at this level of spatial resolution than at the county or state level. Census tracts are small, relatively permanent statistical subdivisions of a county, typically having between 2,500 and 8,000 residents. Census tracts do not cross county boundaries. Their areas vary widely depending on the density of settlement. Census tracts tend to be small in densely populated areas but can be very large in sparsely populated areas. Within census tracts are census blocks, which are areas bounded by visible or virtual features, such as streets, streams, city, or town boundaries. Census blocks are typically small in area; for example, in an urban area, a census block might correspond to a block bounded by city streets. In remote areas, however, census blocks might be large and irregular, comprising many square miles.

Air concentrations are estimated in NATA at various levels of resolution depending upon the source type modeled. Secondary formation, fires, and biogenics (modeled in the CONUS) are at 12-km grid-cell resolution. Other sources use census-block resolution, though the emissions for some sources are at the tract level—these tract-level emissions originate from even broader geographic scales (county and national level) and are less certain at these finer geographies as discussed below. For a given source type and modeling approach, variation in ambient air concentrations within a grid cell or census block is not explicitly modeled. For estimates at the block level, a representative ambient air concentration is estimated for a single location near the center of the block (i.e., the centroid, which is typically, but not always, the geographic center of the block chosen by the U.S. Census Bureau as a reference point). EPA then averages ambient concentrations estimated at the block level for the encompassing census tract, with concentration and risk results reported at the tract level. Assessment results do not reflect variations in the susceptibility of people within a census tract because the focus is to compare typical exposures and risks in different tracts. As a result, individual exposures or risks might differ by as much as a factor of 10 in either direction. Exposure or risk determined in NATA should be considered as representative of the geographic area where an individual lives, but not necessarily be considered as that individual's personal risk.

Thus, the results of the NATA analysis do not allow for a comparison of ambient air concentrations, exposures, or risks between two individuals. They do, however, enable the user to understand the variation in typical values for these quantities among counties or states and to a lesser degree among census tracts. For an individual, however, the values might differ from the typical value for the county or state if that individual lives in a part of the geographic area that has a higher or lower than typical value, has an activity pattern that causes a higher or lower exposure than is typical, or is more (or less) susceptible than a "typical" person used in this assessment.

For the purposes of estimating and reporting risk, EPA assumes that individuals within a census tract have the same exposure and risk. This assumption allows the examination of the variation in individual exposure among census tracts, but it does not allow the examination of the variation within a census tract. Activity patterns are included for each of six cohorts defined by age. Even within a receptor population, some variability in activity patterns among individuals is considered. Differences in susceptibility, however, are not included in NATA. EPA took this approach for NATA for two primary reasons:

- **An overall purpose of NATA is to examine broad differences driven by geography.** NATA considers only geographic differences in pollutant concentration, exposure, and risk. The goal is to understand how these three factors differ among people living in different geographic areas. EPA assesses these differences, as mentioned above, by tracking differences in air concentration in different census tracts, producing differences in the typical pollutant concentrations, exposures, and risks in different tracts. Differences in susceptibility, however, can produce differences in risk between two individuals in the same census tract, and reporting on these differences is not a purpose of NATA.
- **The variability in susceptibility is difficult to model at the national scale.** Very limited information is available on differences in susceptibility among individuals. Even if EPA were to choose to calculate and report differences among individuals in a census tract, scientifically reliable information necessary to produce these calculations is not available for many of the pollutants. Given current information, estimating variability in the rates at which people breathe air might be possible, but this variability is only a small component of the overall variation in susceptibility. EPA therefore has chosen not to incorporate this source of variation between individuals.

Taking into consideration these limitations, EPA elected to incorporate differences in emissions and meteorology (resulting in differences in ambient air concentration) and differences in location of typical individuals (resulting in differences in exposure) among census tracts. Variation in activity patterns for different age groups is reflected in the assessments to the degree than the age of residents varies by location. Variability in susceptibility is not included for the reasons given above. Temporal variation in inputs is addressed in the development of time-weighted averages of emissions characteristics, meteorological conditions, and ECs. Temporal variation in the estimated ambient air concentrations, however, is not reflected in the results (only time-weighted annual averages are presented).

7.2.3 How Variability Affects Interpretation of NATA Results

The NATA analysis illustrates how ambient air concentration, exposure, and risk vary throughout the United States. The assessment does not focus on the variation in exposure and risk among individuals. It focuses on variation among well-defined geographic areas, such as counties or states, based on calculations of ambient air concentration, exposure, and risk in various census tracts. To a lesser degree, variation among demographic groups is also addressed by NATA, in that differences in activity patterns are taken into account in modeling ECs using HAPEM. Risk results, however, are not presented separately for individual demographic groups.

The information contained in the maps, charts, and tables produced in NATA display predictions of cancer risk and noncancer hazard. Cancer risk results include statements such as:

> "X percent of the census tracts in a given area are characterized by a typical lifetime excess cancer risk of less than R."

For this statement, if X is 25 percent and R is 1-in-1 million, the result would be:

> "25 percent of the census tracts are characterized by a typical risk of less than 1-in-1 million."

This statement does not necessarily mean that 25 percent of individuals in the specified area have a cancer risk of less than 1-in-1 million. Some people in these census tracts would be expected to have a risk above 1-in-1 million. Although an individual might live in a census tract where the typical or average risk is less than 1-in-1 million, that individual might live nearer the source than the average person in the census tract, or might have an activity pattern that leads to greater exposure, or might be more susceptible. All these factors could cause that individual to experience a risk above the typical value for that census tract. Conversely, the individual also could have a lower risk by living farther from the source, or having an activity pattern that produces lower exposures, or being less susceptible.

The important point to remember when interpreting the maps and charts of the NATA analysis is that they show variation among values of ambient air concentration, exposure, or risk in census tracts or larger areas such as counties. This presentation allows for the identification of geographic regions (counties or states) where these values are higher or lower than the aggregated national average for all census tracts. It does not allow for the identification of individuals who have higher or lower values of ambient air concentration, exposure, or risk. Nevertheless, individuals with a high risk are more likely to be located in geographic regions characterized by a high risk than in those geographic regions characterized by a low risk. The same can be said for exposure (i.e., individuals with a high exposure are more likely to be found in geographic regions characterized by high exposure than in those regions characterized by low exposure).

7.3 How NATA Addresses Uncertainty

No scientific statement (in risk assessment or other areas of science) can be made with complete confidence. Risk estimates are always uncertain to some degree due to issues such as those discussed below. To maintain transparency and openness in the presentation of risk results, the party conducting a risk assessment must explain these uncertainties and how these uncertainties increase or decrease confidence. The NATA analysis produces statements about variability in ambient air concentrations, exposures, and risks across geographic regions for typical individuals, as described in Section 7.2. In this section, the discussion of uncertainty is intended to address the confidence with which these statements regarding variability can be made. Of importance to note is that uncertainty does not prevent EPA from making a statement of risk, nor does it prevent EPA from taking reasonable actions. Uncertainty does require, however, that the nature of the uncertainty, and the implications for decisions, be understood so the degree of support for the statement can be correctly and properly interpreted.

7.3.1 Components of Uncertainty

Uncertainty arises from a variety of sources. To understand the sources of uncertainty affecting a risk assessment, considering the process by which a study such as NATA is performed is instructive, as described in the following sections.

Problem formulation. The problem to be addressed must first be defined. For example, a question that might help define the problem could include, "Is the occurrence of adverse human health effects correlated with emissions from industrial facilities?" What the study is intended to address and how the results will be used should be clear at the outset. This initial step in the analysis introduces problem-formulation uncertainty. The purpose of NATA is described in Section 1 of this document, where the question addressed in the assessment is defined as precisely as possible (e.g., that the study is limited to estimates of health effects in human populations), along with information about the limitations of the assessment. The issue of problem-formulation uncertainty is not considered further in this document.

Defining the analysis components. This step describes what can influence the answer to the problem. In NATA, the multiple influences include emissions from a variety of sources (e.g., mobile, stationary, biogenic); atmospheric dispersion and chemistry; activity patterns for different cohorts; UREs and RfCs; and other considerations. Where the science is poorly developed, the factors that must be included might not be clear. Resources also might be limited, making the inclusion of all factors in the study infeasible. This step in the analysis, which results in the conceptual model for the assessment, introduces conceptual uncertainty. This issue is also addressed in the discussion of the limitations of NATA in Section 1, where the aspects of the problem that are (and are not) included in the study are addressed (e.g., that the study addresses inhalation of air toxics only). The issue of conceptual uncertainty is not considered further here.

Selecting models. All risk assessments use models. The NATA analysis uses a series of mathematical models. Models are used in NATA to produce the emissions inventory; to calculate ambient air concentration; to calculate exposure; and to calculate risk (for cancer and noncancer effects). All scientific models involve uncertainties because a model reduces a (potentially very complex) set of chemical, biological, physical, social, or other processes to manageable algorithms that can be used to perform calculations and make forecasts. The simplifications that are inherent in the development of a model introduce uncertainties.

Typically, more than one model is available for application to a problem and those models can produce different results. Thus, uncertainty is introduced as to which model, and which model results, should be used. As a simple example, NATA uses a linear statistical model to relate EC and cancer risk: cancer risk equals the exposure (air concentration) multiplied by a URE. Uncertainty analysis involves asking a series of questions: Are we certain this linear relationship is correct? Could the relationship be quadratic (i.e., risk equals exposure multiplied by the square of the dose)? Could the relationship have a threshold (i.e., no risk is apparent until the exposure becomes sufficiently large)? What are the implications for estimates of risk if these different models are used? What are the implications for decisions if a clear choice among the models cannot be made?

This step in the analysis introduces model uncertainty. Judging model uncertainty can be both quantitative and qualitative. Qualitative issues involve the scientific plausibility of the model. Does the model include all important processes? Does it explain the phenomenon (e.g., atmospheric dispersion) well? Is the model well accepted in the scientific community—has it passed critical tests and been subject to rigorous peer review?

Quantitative issues involve comparing model results against sets of data (although this also involves issues of parameter uncertainty discussed in the next bullet). Does the model generally predict these data accurately? Are the predictions accurate to within a factor of 2; a factor of 4? What is the effect of any approximation methods used in the model?

Applying models. The models used in the NATA analysis require parameter inputs such as emission rates, stack heights, fractions of time spent indoors, and UREs. Although models describe general relationships among properties of the real world (e.g., the linear relationship between exposure and cancer risk), parameters quantify these properties for specific cases (e.g., the numerical value of the URE for benzene). Parameters provide the numbers needed in the models. Various databases are available from which these parameters can be estimated, and the methods used to collect the data and to compile the databases introduce uncertainties. All of these factors introduce parameter uncertainty.

Although parameter uncertainty has both quantitative and qualitative aspects, common practice is to characterize this source of uncertainty quantitatively, with some qualitative caveats. For example, parameter uncertainty might be characterized by a confidence interval, which states that the true value of the parameter (such as the stack height for a facility) probably lies somewhere between 40 and 60 meters or that the stack height is "known to be within" a factor of 1.2, or that the stack height is "accurate to within" 20 percent. Attached to this quantitative characterization of uncertainty will be a qualitative caveat such as "the estimate of this uncertainty is based on measurements made in 1990 at facilities similar to the one considered in this study, but a change in the design of stacks might have been made since 1990." This qualitative statement provides some idea of the confidence with which the quantitative assessment of uncertainty can be applied.

7.3.2 Components of Uncertainty Included in NATA

For this discussion, the uncertainties in NATA have been divided into three sources, based on the three steps leading from the estimate of emissions to the calculations of risk. Uncertainty in ambient air concentrations is due to uncertainty in the emissions estimates and in the air quality models. Uncertainty in exposure is due to uncertainty in the activity patterns, the locations of individuals within a census tract, and the microenvironmental concentrations as reflected in the exposure model. Finally, uncertainty in risk is due to uncertainty in the shape of the relationship between exposure and effects, the URE, and the RfC. These three sources of uncertainty are discussed below.

NATA Components that Include Uncertainty

- Ambient concentrations
- Exposure estimates
- Risk estimates

Ambient air concentration. Considering first the predictions of ambient air concentration, the specific sources of uncertainty derive from the parameters for the following: emissions, the stack, particle sizes and reactivity, chemical speciation, terrain, boundary conditions, background concentration, meteorology, and model equations. These sources of uncertainty are discussed briefly in this section.

Emissions parameters, including emission rates and locations of sources, are taken from the NEI database, which is a composite of estimates produced by state and local regulatory agencies, industry, and EPA. Some of these data were further modified during the NATA review. The quality of specific emissions rates and locations in the NEI and resultant NATA emissions (e.g., industrial emissions from a specific census tract) has not been fully assessed, although reviews have been conducted. Some of the parameter values could be out of date, errors might have been introduced in transcribing raw data to a computer file, and other data-quality issues might be present. Emission estimates use a variety of methods such as emission factors, material balances, engineering judgement and source testing. Some release point locations use an average facility location instead of the location of each specific unit within the facility. Release point parameters may be defaulted for some situations. Fugitive release parameters are not required and are defaulted where missing. In addition, TRI data does not provide release point parameters other than identifying sources as "stack" or "fugitive"; the release parameters used historical defaults from previous inventories or new defaults.

Uncertainty also is inherent in the emission models used to develop inventory estimates. For example, county-level air toxic emissions from nonroad equipment are estimated by applying fractions of toxic total hydrocarbons to estimates of county-level hydrocarbons for gaseous air toxics and fractions of toxic particulate matter to estimates of county-level particulate matter for PAHs; emission factors based on milligrams per mile are used for metals. The toxic fractions are derived from speciation data, based on limited testing of a few equipment types. The estimates of county-level total organic gases and particulates are derived from the EPA NONROAD model. In the NONROAD model, uncertainties are associated with emission factors, activity, and spatial-allocation surrogates. National-level emissions in NONROAD are allocated to the county level using surrogates, such as construction costs (to allocate emissions of construction equipment) and employees in manufacturing (to allocate industrial equipment). Availability of more specific local data on equipment populations and usage will result in more accurate inventory estimates. For mobile and nonpoint sources, population is used to allocate vehicle miles traveled from state or metropolitan statistical area to county, which is a source of considerable uncertainty.

For mobile and nonpoint sources, the emissions rates are typically allocated from the county level to census-tract levels through a surrogate such as population or land use. This allocation introduces additional uncertainty because the data on the surrogates also have uncertainty, and the correlations between the surrogates and the emissions are imperfect.

The health effects of a pollutant depend on its chemical form when inhaled. For many sources, the NEI database does not include information on **chemical speciation of the pollutants** of interest, but instead contains the total rate of pollutant emitted in all its forms. Assumptions about chemical speciation are made based on values estimated to be representative at such sources, taking into account information on source type, typical feedstock materials, knowledge of the process involved, or other relevant factors. Any one source, however, might actually have different values than the ones assumed.

The dispersion, or movement, of pollutants in the atmosphere is influenced by the topography of the area surrounding a source, which is characterized by **terrain parameters**. Although the CMAQ model estimates include consideration of topography, the HEM-3 model estimates as implemented for NATA do not in all cases. The HEM-3 model estimates for point sources include consideration of topography, but the estimates for the emissions sources modeled as census tract area sources do not because considering topography in the model requires a single source elevation, which is not always possible for large census tracts. Not accounting for terrain introduces uncertainty into predictions of ambient air concentrations, particularly in areas with hills or mountains.

Another source of uncertainty in the modeling of ambient air concentrations is the values used for the boundary conditions used in CMAQ and **background concentration estimates** that are added to AERMOD concentrations from the non-CMAQ HAPs. These sources might include, for example, contributions from long-range transport of compounds from other counties and states. For more details on background concentrations, refer to the discussion in Section 3.

The representation of **meteorological parameters** in the CMAQ model is advanced, as the parameters are derived using WRF. HEM-3 requires less complex representation of meteorological parameters, primarily the direction and speed of airflow and the stability of the atmosphere (which affects how high gases rise once they are emitted). For HEM-3, NATA uses meteorological data from the nearest available monitoring station or grid cell. Uncertainties arise from the fact that the data typically are not measured at the precise location of a given source and sometimes are not for the same year, and therefore might not represent the meteorological conditions accurately.

The **model equations** used in the air quality models represent another source of uncertainty. The version of HEM-3 used for NATA uses the Gaussian equations implemented in the AERMOD computer model that has been studied extensively. The CMAQ model is more complex in its treatment of pollutant dispersion and atmospheric dynamics; nevertheless, many assumptions underlie its Eulerian approach to dispersion, which are outlined further in the science documentation for the CMAQ model.

While the hybrid approach of combining the CMAQ and HEM-3 models results in improved treatment of chemistry and transport, there are uncertainties in the implementation. The approach requires consistent emissions and meteorological inputs to be used in both models. While emissions were as consistent as possible, some simplifications were necessary. CMAQ was not re-run after all of the emissions changes made during the NATA review; instead, the CMAQ results were adjusted based on HEM-3 adjustments, which could have resulted in some uncertainty. Differences also existed in the spatial and temporal treatment of the emissions. Emissions were allocated from county to tract for HEM-3 and from county to grid cell for CMAQ. While the same underlying data were used for the allocation, there may have been differences (introduced by the irregularly shaped census boundaries) in developing the county-to-tract surrogate fractions and in simplifying the tract boundaries for modeling tract-level emissions. The temporal allocation used in HEM-3 was not exactly the same as in CMAQ for the county-level sources, though average profiles based on the CMAQ temporal approach were developed for use in HEM-3. The HEM-3 meteorology data used the MMIF at every fourth grid cell (as opposed to every grid cell); additionally, the HEM-3 meteorology data also used the non-gridded National Weather Service station data, which were not used in CMAQ. In addition to inconsistencies in model inputs, the hybrid approach uses a HEM-3 grid-cell average for normalizing the individual HEM-3 concentrations within the grid cell. The HEM-3 surface values are less representative of the true HEM-3 average in grid cells where there are fewer census-block receptors and where the block receptors are clustered unevenly within the grid cells.

To help characterize the aggregate uncertainty of the predictions of the air quality models, EPA compared modeled concentrations to available monitoring data on ambient air quality. For each monitor-pollutant combination, EPA compared the predicted annual-average concentrations at the monitor location to the sampled annual-average concentrations. These comparisons showed reasonably good agreement. Measured concentrations were taken from EPA's Ambient Monitoring Archive which includes National Air Toxics Trends Stations and state and local monitors reported to the Air Quality System. For the 2011 NATA, the exact locations of the monitors were used for the model-to-monitor comparison, an approach that increases accuracy over previous assessments. For more details about the model-to-monitor analyses for previous assessments, see Comparison of 1996 ASPEN Modeling System Results to Monitored Concentrations (EPA 2002c), Comparison of 1999 Model-Predicted Concentrations to Monitored Data (EPA 2006b), Comparison of 2002 Model-Predicted Concentrations to Monitored Data (EPA 2009), and Comparison of 2005 Model-Predicted Concentrations to Monitored Data (EPA 2010b).

Discrepancies between model predictions and concentration measurements can be attributed to five sources of uncertainty:

- emission characterization (e.g., specification of source location, emission rates, and release characterization);

- meteorological characterization (e.g., representativeness);
- model formulation and methodology (e.g., characterization of dispersion, plume rise, deposition, chemical reactivity);
- monitoring; and
- boundary conditions/background concentrations.

Underestimates for some pollutants could be a result of the following:

- The NEI might be missing specific emission sources (some of the emissions parameters are missing for many of the sources in the NEI).
- The emission rates could be underestimated or overestimated due to emission-estimation techniques and/or spatial allocation of national estimates to county, and county estimates to tracts.
- The accuracy of the monitor averages is uncertain; the monitors, in turn, have their own sources of uncertainty. Sampling and analytical uncertainty, measurement bias, and temporal variation all can cause the ambient concentrations to be inaccurate or imprecise representations of the true atmospheric averages.
- Model-to-model spatial comparisons are imprecise. The results suggest that the model estimates are uncertain on a local scale (i.e., at the census-tract level). EPA believes that the model estimates are more reliably interpreted as being a value likely to be found within 30 km of the census-tract location.

Exposure. Sources of uncertainty in the relationship between ambient air concentrations and ECs include those associated with microenvironmental factors and activity patterns. HAPEM calculates the EC in various microenvironments (e.g., indoors at home, in a car) based on inputs of predicted ambient air concentrations and **microenvironmental factors**. The factors are characterized as probability distributions to reflect the variability found in air-toxics measurements more fully. For many air toxics, the measurement studies needed to estimate microenvironmental factors are not available, so the values used are based on measurement studies of similar compounds in similar situations. This practice introduces uncertainty into the estimation of ECs for such compounds. In addition, even for air toxics with measurement studies, the estimated microenvironmental factors have some uncertainty because the number of such studies is limited. Furthermore, the uniform application of the microenvironmental factors to all census tracts introduces uncertainty by not accounting for possible geographic differences among tracts (e.g., different window-opening behavior, different levels of building integrity).

The **activity-pattern sequences for individuals** used in HAPEM are based on CHAD. As explained in Section 4.3.3, the algorithms in HAPEM consider the variability in activity patterns among individuals within a cohort-tract combination, largely by addressing correlation between subsequent activity patterns assumed to occur for each cohort-tract combination. The representativeness of the daily diaries in CHAD is uncertain because they are a compilation of many studies, including some that are not recent and some for which the data are based on non-random sampling. How well the model algorithms represent actual daily autocorrelation between types of activity also is uncertain. This latter issue, however, pertains only to the variability of the ECs across the demographic group and not the median EC, which is the concentration reported by NATA.

The **commuting data** used in HAPEM are based on an EPA analysis of information from a special study by the U.S. Census. HAPEM uses this information, reflecting 2010 data, in coordination with the activity-pattern data to place an individual either in the home tract or the work tract at each time step. These data introduce some uncertainty because they simplify commuting patterns to a pair of home and work census tracts and might not reflect certain details of some commutes (e.g., the additional census tracts encountered by commuters who travel to non-adjacent tracts; more complex commuting patterns that are not point to point). An additional important consideration is that the commuting-pattern data included in HAPEM do not account for the movement of school-age children who travel (or commute) to a school located outside the tracts in which they reside.

Risk. Concerning the predictions of risk, the specific sources of uncertainty in dose-response relationships (in addition to those considered for ambient air concentration and exposure) are hazard identification, dose-response models for carcinogens, UREs, and RfCs.

One component of predicting risk is **hazard identification**. Cancer-risk estimates are based on the assumption that a compound either is a carcinogen or produces a noncancer effect. This judgment is based on the results of a hazard-identification stage in which the evidence that an air toxic produces either cancer or a noncancer effect is assessed. Because the evidence for either judgment is never unequivocal, a compound labeled as a carcinogen or one deemed to produce noncancer effects, in fact, might produce no such effect in humans. This possibility introduces uncertainty into the calculation of risk because the risk, in fact, could be zero. As the evidence for the original conclusion (i.e., that the compound produces the effect) increases, this uncertainty decreases.

Cancer-risk estimates are based on the assumption that the relationship between exposure and probability of cancer is linear. In other words, the probability of developing cancer is assumed proportional to the exposure (equal to the exposure multiplied by a URE). This type of **dose-response model** is used routinely in regulatory risk assessment because it is believed to be conservative; that is, if the model is incorrect, it is more likely to lead to an overestimate of the risk than to an underestimate. Other scientifically valid, biologically based models are available, which produce estimates of cancer risk that differ from those obtained from the linear model. Uncertainty in risk estimates therefore, is, introduced by the inability to justify completely the use of one model or the other (because each model has some scientific support). An essential consideration is that this uncertainty is, to some extent, one-sided. In other words, conservatism when uncertainty exists allows more confidence in the conclusion that the true risk is *less than that predicted* than in the conclusion that the risk is *greater than that predicted*.

URE parameters have associated uncertainty. In some cases, the UREs are based on maximum-likelihood estimates of the slope of the dose-response relationship derived from reliable data. In other cases, the UREs are based on "upper-bound" estimates (i.e., the slope is not the best estimate, but is a conservative value that is likely to lead to overestimates of risk) derived from less reliable data. For some compounds, the UREs are derived from human-exposure studies, but for others they are from animal exposures. These considerations introduce uncertainty into the URE values, and the amount of uncertainty varies among pollutants.

Another source of uncertainty in estimating risk derives from the values chosen for the **RfC parameters** used to calculate an HQ for noncancer health risk. The RfC, which (like the URE) is based on limited information, is uncertain, and as a result, the value of HQ is uncertain. As is the case for UREs, the uncertainty in the RfC is generally one-sided and the risk is unlikely to be greater than predicted.

7.4 Summary of Limitations in NATA

EPA developed this assessment to inform both national and more localized efforts to collect information and characterize or reduce air-toxics emissions (e.g., to prioritize pollutants or geographic areas of interest for monitoring and community assessments). As described above, many of the elements in the assessment process for NATA, as in other assessments that derive results from environmental data and modeling of environmental data, are characterized by uncertainty and variability. Because of this, EPA suggests exercising caution when using the results of these assessments, as the overall quality and uncertainty of each assessment vary from location to location and from pollutant to pollutant. In many cases assessments that are more localized, incorporating appropriately scaled local monitoring and modeling, could be necessary to better characterize local-level risk.

Recognizing the specific limitations in NATA results is critical to their proper interpretation and utility, including that the results:

- apply to geographic areas, not specific locations,
- do not include comprehensive impacts from sources in Canada or Mexico,
- are restricted to the year to which the assessment pertains (because the assessment uses emissions data from that year),
- do not reflect exposures and risk from all compounds,
- do not reflect all pathways of exposure,
- reflect only compounds released into the outdoor air,
- do not fully capture variations in background ambient air concentrations,
- might underestimate or overestimate ambient air concentrations for some compounds due to spatial uncertainties,
- are based on default, or simplifying, assumptions where data are missing or of poor quality, and
- might not accurately capture sources that have episodic emissions, and contain uncertainty.

The results apply to geographic areas, not specific locations. The assessment focuses on variations in air concentration, exposure, and risk among geographic areas such as census tracts, counties, and states. All questions asked, therefore, must focus on the variations among different areas. They cannot be used to identify "hot spots" where the air concentration, exposure, or risk might be significantly higher than other locations. Furthermore, this type of modeling assessment cannot address the kinds of questions an epidemiology study might, such as the relationship between asthma or cancer risk or proximity of residences to point sources, roadways, and other sources of pollutant emissions.

The results do not include comprehensive impacts from sources in Canada or Mexico. The NATA results for states that border these countries do not thoroughly reflect these potentially significant sources of transported emissions.

The results apply to groups, not to specific individuals. Within a census tract, all individuals are assigned the same ambient air concentration, which is chosen to represent a typical ambient air concentration. Similarly, the exposure assessment uses activity patterns that do not fully reflect variations among individuals. As a result, the exposures and risks in a census tract should be interpreted as typical values rather than as means, medians, or some other statistical average. The values are likely to be in the midrange of values for all individuals in the census tract.

The results for the 2011 NATA are restricted to 2011 because the assessment used emissions data from 2011. Also, the assumption regarding emissions in the assessment is that the levels remain constant throughout one's lifetime (the emissions are not today's levels nor are they projected levels). Emissions continue to decrease, however, as (1) mobile-source regulations are phased in over time, (2) EPA-issued air-toxics regulations for major industrial sources reach compliance due dates, (3) state and industry initiatives to reduce air pollutants continue, and (4) some facilities are closed or have made process changes or other changes that have significantly reduced their emissions since 2011.

The results do not reflect exposures and risk from all compounds. Only 138 of the 181 air toxics (i.e., 180 CAA HAPs plus diesel PM) modeled in NATA have dose-response values. The remaining 43 air toxics do not and therefore are not considered in the aggregate cancer risk or target-organ-specific hazard indices. Of particular significance is that the assessment does not quantify cancer risk from diesel PM, although EPA has concluded that the general population is exposed to levels close to or overlapping with apparent levels that have been linked to increased cancer risk in epidemiology studies. Currently, a URE for diesel PM has not yet been derived; therefore, a quantitative estimate of the cancer risks has not been included in the 2011 NATA. An IRIS RfC for diesel PM has allowed a quantitative estimate of the noncancer effects.

The results do not reflect all pathways of exposure. The assessment includes only risks from direct inhalation of the emitted pollutants. It does not consider pollutants that might then deposit onto soil and into water and food, and therefore enter the body through ingestion or skin contact. Consideration of these routes of exposure could increase estimates of exposure and risk.

The assessment results reflect only compounds released into the outdoor air. The assessment does not include exposure to pollutants produced indoors, such as from stoves or out-gassing from building materials, or evaporative benzene emissions from cars in attached garages. For some compounds such as formaldehyde, these indoor sources can contribute significantly to the total exposure for an individual, even if only inhalation exposures are considered. In addition, the assessment does not consider pollutants released directly to water and soil. It does take into account transformation of one pollutant into another (i.e., secondary formation) in the atmosphere.

The assessment does not utilize CMAQ in all areas (i.e., not in Alaska, Hawaii, Puerto Rico, and the U.S. Virgin Islands) and therefore does not estimate fires, biogenics, and secondary formation based on location-specific data in these areas. It also does not utilize CMAQ for all pollutants and hence may not appropriately estimate the long-range transport for these non-CMAQ pollutants. For pollutants not estimated in CMAQ, the assessment uses background ambient air concentrations that are based on remote concentration estimates, but these would not account for variations due to the regional transport of these pollutants.

The assessment might underestimate or overestimate ambient air concentrations for some compounds in some locations due to spatial uncertainty in mobile and nonpoint emissions, which are more uncertain at finer geographic scales.

The assessment uses default, or simplifying, assumptions where data are missing or of poor quality. Data for some variables used in the modeling for emissions and dispersion of pollutants (such as stack height and facility location) are not always available or are flawed. In such instances, these values are replaced by default assumptions. For example, a stack height for a facility might be set equal to stack heights at comparable facilities or the location of the release points within a facility might be placed at the center of the facility. These substitutions introduce uncertainty into the final predictions of ambient concentration, exposure, and risk.

The assessment might not accurately capture sources that have episodic emissions. **Some facilities might experience short-term (a few days or weeks) deviations from their typical emissions patterns, such**

as during startups, shutdowns, malfunctions, and upsets. NATA modeling assumes that emission rates are uniform throughout the year.

8 REFERENCES

Akhtar, F., Henderson, B., Appel, W., Napelenok, S., Hutzell, B., Pye, H., and Foley, K. 2012. Multiyear Boundary Conditions for CMAQ 5.0 from GEOS-Chem with Secondary Organic Aerosol Extensions. 11th Annual Community Modeling and Analysis System Conference, Chapel Hill, NC, October 2012.

ATSDR (Agency for Toxic Substances and Disease Registry). 2015. Toxic Substances Portal Minimal Risk Levels (MRLs) for Hazardous Substances. Available online at http://www.atsdr.cdc.gov/mrls/index.asp. Last updated 28 October 2015. Last accessed 10 December 2015.

Appel, K.W., Gilliam, R.C., Davis, N., Zubrow, A., and Howard, S.C. 2011. Overview of the Atmospheric Model Evaluation Tool (AMET) v1.1 for Evaluating Meteorological and Air Quality Models. *Environ. Modell. Softw.*, 26(4): 434–443.

Attfield, M.D., Schleiff, P.L., Lubin, J.H., Blair, A., Stewart, P.A., Vermeulen, R., Coble, J.B., and Silverman, D.T. 2012. The Diesel Exhaust in Miners Study: A Cohort Mortality Study with Emphasis on Lung Cancer. *Journal of the National Cancer Institute*, 104: 1–15. Available online at http://jnci.oxfordjournals.org/content/104/11/855. Last accessed 10 December 2015.

Brown, N., Allen, D., Amar, P., Kallos, G., McNider, R., Russell,, A., and Stockwell, W. 2011. Final Report: Fourth Peer Review of the CMAQ Model. EPA/ORD/NERL, Research Triangle Park, NC. Available online at http://cfpub.epa.gov/si/si_public_file_download.cfm?p_download_id=525232. Last accessed 11 December 2015.

Byun, D., and Schere, K.L. 2006. Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. *Applied Mechanics Reviews*, 59(2): 51–77.

Caldwell, J.C., Woodruff, T.J., Morello-Frosch, R., and Axelrad, D.A. 1998. Application of Health Information to Hazardous Air Pollutants Modeled in EPA's Cumulative Exposure Project. *Toxicology and Industrial Health*, 14(3): 429–454.

Cimorelli, A.J., Perry, S.G., Venkatram, A., Weil, J.C., Paine, R.J., Wilson, R.B., Lee, R.F., Peters, W.D., and Brode, R.W. 2005. AERMOD: A Dispersion Model for Industrial Source Applications. Part I: General Model Formulation and Boundary Layer Characterization. *Journal of Applied Meteorology*, 44: 682–693.

Cook, R., Phillips, S., Houyoux, M., Dolwick, P., Mason, R., Yanca, C., Zawacki, M., Davidson, K., Michaels, H., Harvey, C., Somers, J., and Luecken, D. 2011. Air Quality Impacts of Increased Use of Ethanol under the United States' Energy Independence and Security Act. *Atmospheric Environment*, 45: 7714–7724.

EPA (U.S. Environmental Protection Agency). 1986. Guidelines for Mutagenicity Risk Assessment. EPA/630/R-98/003. EPA, Washington, DC. Available online at http://www2.epa.gov/risk/guidelines-mutagenicity-risk assessment. Last accessed 10 December 2015.

EPA. 1991. Guidelines for Developmental Toxicity Risk Assessment. EPA/600/R-91/001. EPA, Washington, DC. Available online at http://www2.epa.gov/risk/guidelines-developmental-toxicity-risk assessment. Last accessed 26 October 2015.

EPA. 1994. Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry. EPA/600/8-90/066F. EPA Office of Research and Development (ORD), Washington, DC. Available online at http://www2.epa.gov/risk/methods-derivation-inhalation-reference-concentrations-and-application-inhalation-dosimetry. Last accessed 27 October 2015.

EPA. 1996. Guidelines for Reproductive Toxicity Risk Assessment. EPA/630/R-96/009. EPA, Washington, DC. Available online at http://www2.epa.gov/risk/guidelines-reproductive-toxicity-risk assessment. Last accessed 10 December 2015.

EPA. 1998. Guidelines for Neurotoxicity Risk Assessment. EPA/630/R-97/0. EPA, Washington, DC. Available online at http://www2.epa.gov/risk/guidelines-neurotoxicity-risk assessment. Last accessed 10 December 2015.

EPA. 2000. Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures. EPA/630/R-00/002. Risk Assessment Forum, Washington, DC. Available online at http://ofmpub.epa.gov/eims/eimscomm.getfile?p_download_id=4486. Last accessed 8 December 2015.

EPA. 2001a. NATA—Evaluating the National-scale Air Toxics Assessment 1996 Data – An SAB Advisory. EPA/SAB/EC/ADV-02/001. Science Advisory Board, Washington, DC. Available online at http://archive.epa.gov/airtoxics/nata/web/pdf/sabrept1201.pdf. Last accessed 29 November 2015.

EPA. 2001b. National-scale Air Toxics Assessment for 1996. Draft for EPA Science Advisory Board Review: January 18, 2001. EPA-453/R-01-003. EPA Office of Air Quality Planning and Standards (OAQPS), Research Triangle Park, NC. Available online at http://archive.epa.gov/airtoxics/nata/web/html/sabrev.html. Last accessed 29 November 2015.

EPA. 2002a. Health Assessment Document for Diesel Engine Exhaust. EPA/600/8-90/057F. EPA ORD/ National Center for Environmental Assessment (NCEA), Washington, DC. Available online at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=29060. Last accessed 4 December 2015.

EPA. 2002b. The HAPEM User’s Guide Hazardous Air Pollutant Exposure Model, Version 4. EPA OAQPS, Research Triangle Park, NC. Available online at http://archive.epa.gov/airtoxics/nata/web/zip/hapem4guide4.zip. Last accessed 29 November 2015.

EPA. 2002c. Comparison of ASPEN Modeling System Results to Monitored Concentrations. EPA. Available online at http://archive.epa.gov/airtoxics/nata/web/html/mtom_pre.html. Last updated 22 October 2015. Last accessed 30 November 2015.

EPA. 2003. Framework for Cumulative Risk Assessment. EPA/630/P-02/001F. EPA ORD/NCEA, Washington, DC. Available online at http://www2.epa.gov/risk/framework-cumulative-risk assessment. Last accessed 26 October 2015.

EPA. 2004a. Air Toxics Risk Assessment Reference Library. Volume 1: Technical Resource Manual. EPA-453/K-04-001A. EPA OAQPS, Research Triangle Park, NC. Available online at http://www2.epa.gov/sites/production/files/2013-08/documents/volume_1_reflibrary.pdf. Last accessed 2 December 2015.

EPA. 2004b. Air Toxics Risk Assessment Reference Library. Volume 2: Facility-Specific Assessment. EPA-453/K-04-001B. EPA OAQPS, Research Triangle Park, NC. Available online http://www2.epa.gov/sites/production/files/2013-08/documents/volume_2_facilityassess.pdf. Last accessed 2 December 2015.

EPA. 2005a. Guidelines for Carcinogen Risk Assessment. EPA/630/P-03/001F. EPA, Washington, DC. Available online at http://www2.epa.gov/risk/guidelines-carcinogen-risk assessment. Last accessed 26 October 2015.

EPA, 2005b. EPA's National Mobile Inventory Model (NMIM), a Consolidated Emissions Modeling System for MOBILE6 and NONROAD. EPA-420-R-05-024. EPA Office of Transportation and Air Quality (OTAQ) Assessment and Standards Division, Ann Arbor, MI. Available online at http://www.epa.gov/otaq/models/nmim/420r05024.pdf. Last accessed 10 December 2015.

EPA. 2005c. Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens. EPA/630/R-03/003F. EPA, Washington, DC. Available online at http://www2.epa.gov/osa/memoranda-about-implementation-cancer-guidelines-and-accompanying-supplemental-guidance-science. Last accessed 4 December 2015.

EPA. 2005d. The HAPEM User's Guide Hazardous Air Pollutant Exposure Model, Version 5. EPA OAQPS, Research Triangle Park, NC. Available online at http://www2.epa.gov/sites/production/files/2013-08/documents/hapem5_guide.pdf. Last accessed 29 November 2015.

EPA. 2006a. Air Toxics Risk Assessment Reference Library. Volume 3: Community-Scale Assessment. EPA-453/K-06-001C. EPA OAQPS, Research Triangle Park, NC. Available online at http://www2.epa.gov/sites/production/files/2013-08/documents/volume_3_communityassess.pdf. Last accessed 2 December 2015.

EPA. 2006b. Comparison of 1999 Model-Predicted Concentrations to Monitored Data. EPA OAQPS. Available online at http://archive.epa.gov/airtoxics/nata1999/web/html/99compare.html. Last updated 13 September 2015. Last accessed 30 November 2015.

EPA. 2007. The HAPEM User's Guide Hazardous Air Pollutant Exposure Model, Version 6. EPA OAQPS, Research Triangle Park, NC. Available online at http://www2.epa.gov/sites/production/files/2013-08/documents/hapem6_guide.pdf. Last accessed 29 November 2015.

EPA. 2008a. Health Effects Assessment Summary Tables (HEAST). EPA NCEA, Washington, DC. Available online at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=2877. Last updated 21 July 2008. Last accessed 4 December 2015.

EPA. 2008b. The Clean Air Act Amendments of 1990 List of Hazardous Air Pollutants. EPA Office of Air and Radiation (OAR), Washington, DC. Available online at http://www.epa.gov/ttn/atw/orig189.html. Last updated 10 September 2015. Last accessed 10 December 2015.

EPA. 2009. Comparison of 2002 Model-Predicted Concentrations to Monitored Data. EPA OAQPS. Available online at http://archive.epa.gov/nata2002/web/html/compare.html. Last accessed 30 November 2015.

EPA. 2010a. Development of a Relative Potency Factor (RPF) Approach for Polycyclic Aromatic Hydrocarbon (PAH) Mixtures: In Support of Summary Information of the Integrated Risk Information System (IRIS) (External Review Draft). EPA/635/R-08/012A. EPA, Washington, DC. Available online at http://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=194584. Last accessed 19 October 2015.

EPA. 2010b. Results of the 2005 NATA Model-to-Monitor Comparison, Final Report. Prepared by Eastern Research Group for EPA OAQPS. Available online at http://www3.epa.gov/ttn/atw/nata2005/05pdf/nata2005_model2monitor.pdf. Last accessed 30 November 2015.

EPA. 2011. An Overview of Methods for EPA's National-scale Air Toxics Assessment. Prepared by ICF International for EPA OAQPS. Available online at http://www3.epa.gov/ttn/atw/nata2005/05pdf/nata_tmd.pdf. Last accessed 10 December 2015.

EPA. 2014a. Meteorological Model Performance for Annual 2011 WRF v3.4 Simulation. EPA. Available online at http://www.epa.gov/ttn/scram/reports/MET_TSD_2011_final_11-26-14.pdf. Last accessed 10 December 2015.

EPA. 2014b, Profile of the 2011 National Air Emissions Inventory, EPA. Available online at http://www3.epa.gov/ttn/chief/net/lite_finalversion_ver10.pdf. Last accessed 10 December 2015.

EPA. 2014c. Toxicological Review of Trichloroethylene. EPA NCEA, Washington, DC. Available online at http://cfpub.epa.gov/ncea/iris/search/index.cfm?keyword=trichloroethylene. Last updated 12 September 2014. Last accessed 10 December 2015.

EPA. 2014d. Draft User's Manual for The Mesoscale Model Interface Program (MMIF) Version 3.1. Prepared by ENVIRON International Corporation for EPA OAQPS. Available online at http://www3.epa.gov/scram001/models/relat/mmif/MMIFv3.1_Users_Manual.pdf. Last accessed 14 December 2015.

EPA. 2014e. Human Exposure Model (HEM-3) User's Guides. Prepared by EC/R Incorporated for EPA OAQPS. Available online at http://www.epa.gov/fera/human-exposure-model-hem-3-users-guides. Last accessed 24 March 2015.

EPA. 2015a. Consolidated Human Activity Database (CHAD). EPA, Washington, DC. Available online at http://www2.epa.gov/healthresearch/consolidated-human-activity-database-chad-use-human-exposure-and-health-studies-and. Last updated 30 September 2015. Last accessed 17 November 2015.

EPA. 2015b. MOVES 2014a. EPA OTAQ. Available online at http://www3.epa.gov/otaq/models/moves/index.htm. Last updated 1 December 2015. Last accessed 10 December 2015.

EPA. 2015c. The HAPEM User's Guide Hazardous Air Pollutant Exposure Model, Version 7. EPA OAQPS, Research Triangle Park, NC. Available online at http://www.epa.gov/fera/hazardous-air-pollutant-exposure-model-hapem-users-guides. Last accessed 29 November 2015.

EPA. 2015d. What Are the Six Common Air Pollutants? EPA, Washington, DC. Available online at http://www3.epa.gov/airquality/urbanair/. Last updated 18 September 2015. Last accessed 2 December 2015.

EPA. 2015e. Risk Assessment Guidance and Tools. EPA, Washington, DC. Available online at http://www.epa.gov/risk/guidance.htm. Last updated 1 December 2015. Last accessed 2 December 2015.

EPA. 2015f. User's Guide for the AMS/EPA Regulatory Model – AERMOD. EPA-454/B-03-001. Addendum June 2015. EPA, Research Triangle Park, NC.

EPA. 2015g. Community Multiscale Air Quality (CMAQ). EPA, Washington, DC. Available online at http://www.epa.gov/air-research/community-multi-scale-air-quality-cmaq-modeling-system-air-quality-management. Last updated 8 December 2015. Last accessed 10 December 2015.

EPA. 2015h. Clean Air Markets. EPA, Washington, DC. Available online at http://www.epa.gov/airmarkt/. Last updated 17 November 2015. Last accessed 10 December 2015.

EPA. 2015i. Risk and Technology Review. EPA OAR, Washington, DC. Available online at http://www.epa.gov/ttn/atw/rrisk/rtrpg.html. Last updated 1 December 2015. Last accessed 10 December 2015.

EPA. 2015j. Toxics Release Inventory Program. EPA, Washington, DC. Available online at http://www.epa.gov/toxics-release-inventory-tri-program. Last updated 13 November 2015. Last accessed 10 December 2015.

EPA. 2015k. MOVES 2014. EPA OTAQ, Washington, DC. Available online at http://www3.epa.gov/otaq/models/moves/moves-docum.htm. Last updated 5 November 2015. Last accessed 10 December 2015.

EPA. 2015l. Products and Publications Relating to Risk Assessment Produced by the Office of the Science Advisor (OSA). EPA OSA, Washington, DC. Available online at http://www.epa.gov/osa/products-and-publications-relating-risk-assessment-produced-office-science-advisor. Last updated 3 March 2015. Last accessed 8 December 2015.

EPA. 2015m. Evaluation of Prognostic Meteorological Data in AERMOD Applications. EPA-454/R-15-004. EPA OAQPS, Research Triangle Park, NC. Available online at http://www3.epa.gov/ttn/scram/11thmodconf/MMIF_Evaluation_TSD.pdf. Last accessed 14 December 2015.

EPA. 2015n. Overview by Section of CAA. EPA OAR, Washington, DC. Available online at http://www.epa.gov/ttn/atw/overview.html. Last updated 10 September 2015. Last accessed 10 December 2015.

EPA. 2015o. MOBILE Model (onroad vehicles). EPA OTAQ, Washington, DC. Available online at http://www.epa.gov/otaq/mobile.htm. Last updated 5 November 2015. Last accessed 10 December 2015.

EPA. 2015p. National Mobile Inventory Model (NMIM). EPA OTAQ, Washington, DC. Available online at http://www.epa.gov/otaq/nmim.htm. Last updated 5 November 2015. Last accessed 10 December 2015.

EPA. 2015q. NONROAD Model (nonroad engines, equipment, and vehicles). EPA OTAQ, Washington, DC. Available online at http://www.epa.gov/otaq/nonrdmdl.htm. Last updated 5 November 2015. Last accessed 10 December 2015.

FAA (U.S. Federal Aviation Administration). 2015. Airport Data & Contact Information. FSS. Available online at http://www.faa.gov/airports/airport_safety/airportdata_5010/. Last updated 3 November 2015. Last accessed 10 December 2015.

Garshick, E., Laden, F., Hard, J.E., Davis, M.E., Eisen, E.A., and Smith, T.J. 2012. Lung Cancer and Elemental Carbon Exposure in Trucking Industry Workers. *Environmental Health Perspectives*, 120:1301–1306. Available online at http://dx.doi.org/10.1289/ehp.1204989. Last accessed 10 December 2015.

HEI (Health Effects Institute). 2015. Diesel Emissions and Lung Cancer: An Evaluation of Recent Epidemiological Evidence for Quantitative Risk Assessment. HEI, Boston, MA. Available online at http://pubs.healtheffects.org/view.php?id=446. Last updated 24 November 2015. Last accessed 10 December 2015.

Henderson, B.H., Akhtar, F., Pye, H.O.T., Napelenok, S.L., and Hutzell, W.T. 2014. A Database and Tool for Boundary Conditions for Regional Air Quality Modeling: Description and Evaluation. *Geosci. Model Dev.*, 7:339–360.

Houyoux, M.R., Vukovich, J.M., Coats, C.J. Jr., Wheeler, N.M., and Kasibhatla, P.S. 2000. Emission Inventory Development and Processing for the Seasonal Model for Regional Air Quality (SMRAQ) Project. *Journal of Geophysical Research,* 105(D7):9079–9090.

IARC (International Agency for Research on Cancer). 2013. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Volume 105 (2013). Available online at http://monographs.iarc.fr/ENG/Monographs/vol105/index.php. Last accessed 10 December 2015.

Isakov, V., Irwin., J., and Ching, J.K. 2007. Using CMAQ for Exposure Modeling and Characterizing the Sub-grid Variability for Exposure Estimates. *Journal of Applied Meteorology and Climatology,* 46:1354–1371.

NRC (National Research Council). 1983. Risk Assessment in the Federal Government: Managing the Process. Committee on the Institutional Means for Assessments of Risk to Public Health, Commission on Life Sciences. NRC. National Academy Press, Washington, DC.

NRC. 1994. Science and Judgment in Risk Assessment. Committee on Risk Assessment of Hazardous Air Pollutants, Board on Environmental Sciences and Technology, Commission on Life Sciences. NRC. National Academy Press, Washington, DC.

NUATRC (Mickey Leland National Urban Air Toxics Research Center). 2011. Available online at http://www.sph.uth.tmc.edu/mleland/. Last updated 30 December 2011. Last accessed 10 December 2015.

OEHHA (Office of Environmental Health Hazard Assessment, California). 2014. Air Toxicology and Epidemiology. OEHHA, Sacramento, CA. Available online at http://www.oehha.ca.gov/air/allrels.html. Values last updated June 2014. Last accessed 10 December 2015.

OEHHA. 2015. Hot Spots Guidelines. OEHHA, Sacramento, CA. Available online at http://www.oehha.ca.gov/air/hot_spots/index.html. Contents last updated 14 August 2015. Last accessed 10 December 2015.

Otte T.L., and Pleim, J.E. 2010. The Meteorology-Chemistry Interface Processor (MCIP) for the CMAQ Modeling System: Updates through v3.4.1. *Geoscientific Model Development,* 3:243–256.

Russell, A.G., and Dennis, R. 2000. NARSTO Critical Review of Photochemical Models and Modeling. *Atmospheric Environment*, 3:2283–2324.

Seigneur, C., and Dennis, R. 2010. Technical Challenges in Multipollutant Air Quality Management. edited by: Hidy, G.M., Brook, J.R., Demerjian, K.L., Molina, L.T., Pennell, W.T., and Scheffe, R.D. Springer, Dordrecht.

Silverman, D.T., Samanic, C.M., Lubin, J.H., Blair, A.E., Stewart, P.A., Vermeulen, R., Coble, J.B., Rothman, N., Schleiff, P.L., Travis, W.D., Ziegler, R.G., Wacholder, S., and Attfield, M.D. 2012. The Diesel Exhaust in Miners Study: A Nested Case-Control Study of Lung Cancer and Diesel Exhaust. Available online at http://dx.doi.org/10.1093/jnci/djs034. Last accessed 10 December 2015.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X., Wang, W., and Powers, J.G. 2008. A Description of the Advanced Research WRF Version 3. Available online at http://www2.mmm.ucar.edu/wrf/users/pub-doc.html. Last updated 5 December 2014. Last accessed 16 December 2015.

Wesson, K., Fann, N., Morris, M., Fox, T., and Hubbell, B. 2010. A Multi-pollutant, Risk-based Approach to Air Quality Management: Case Study for Detroit. *Air Pollut. Res.*, 1:296–304. doi:10.5094/APR.2010.037.

WHO (World Health Organization). 2015. Complete List of Agents Evaluated and their Classification. WHO IARC, Lyon, France. Available online at http://monographs.iarc.fr/ENG/Classification/index.php. Last updated 26 October 2015. Last accessed 10 December 2015.

Yantosca, B., Sulprizio, M., Yannetti, M., Lundgren, L., and Xu, J. 2015. GEOS-Chem v10-01 Online User's Guide, Atmospheric Chemistry Modeling Group, School of Engineering and Applied Sciences, Harvard University, Cambridge, MA. Available online at http://acmg.seas.harvard.edu/geos/doc/man/. Last accessed 10 December 2015.

Yarwood, G., Rao, S., and Whitten, G. 2005. Updates to the Carbon Bond Mechanism, CB05, Final Report, RT-04-00675. Prepared by ENVIRON International Corporation for EPA. Available online at http://www.camx.com/files/cb05_final_report_120805.aspx. Last accessed 3 December 2015.

This page intentionally left blank.

Appendix A

Glossary

"N"-in-1 million cancer risk:

A risk level of "N"-in-1 million implies a likelihood that up to "N" people, out of one million equally exposed people would contract cancer if exposed continuously (24 hours per day) to the specific concentration over 70 years (an assumed lifetime). This would be in addition to those cancer cases that would normally occur in an unexposed population of one million people. Note that this assessment looks at *lifetime* cancer risks, which should not be confused with or compared to *annual* cancer risk estimates. If you would like to compare an annual cancer risk estimate with the results in this assessment, you would need to multiply that annual estimate by a factor of 70 or alternatively divide the lifetime risk by a factor of 70.

Activity-pattern data:

In an inhalation exposure assessment, activity-pattern data depict both the actual physical activity (including an associated inhalation exertion level); the physical location; and, the time of day the activity takes place (e.g., at midnight, while sleeping at home, jogging in the park at 8 a.m., or driving in a car at 6 p.m.). The Hazardous Air Pollution Model (HAPEM) uses activity-pattern data from EPA's Comprehensive Human Activity Database (CHAD).

AMS/EPA Regulatory Model (AERMOD):

EPA's preferred model for near-field (i.e., within 50 km) simulations of dispersion of emissions. In simulating boundary-layer turbulence, it has the capability to model complex terrain, elevated sources, numerous discrete receptors, and source types ranging from point to line to volume, at hourly resolution.

Air toxics:

Also known as toxic air pollutants or hazardous air pollutants*; those pollutants known to cause or suspected of causing cancer or other serious health problems. Health concerns could be associated with both short- and long-term exposures to these pollutants. Many are known to have respiratory, neurological, immune, or reproductive effects, particularly for more susceptible or sensitive populations such as children. Five important air pollutants are not included in the list of air toxics because the Clean Air Act addresses them separately as "criteria pollutants." These are particulate matter (PM), nitrogen oxides (NO_x), sulfur oxides (SO_x), ozone, and carbon monoxide. Lead is both a criteria pollutant and an air toxic. Criteria pollutants are not addressed in NATA.

*Diesel particulate matter is not a hazardous air pollutant but is included in the NATA air toxics.

Ambient:

Surrounding, as in the surrounding environment. In NATA assessments, ambient air refers to the outdoor air surrounding a person through which pollutants can be carried. Therefore, the ambient concentrations estimated by NATA are those concentrations estimated in the outdoor environment. NATA also estimates exposure concentrations that result from an individual's movement through various microenvironments, including the indoor environment.

Area and other sources:

Include sources that generally have lower emissions on an individual basis than "major sources" and are often too small or ubiquitous to be inventoried as individual sources. "Area sources" include facilities that have air toxics emissions below the major source threshold as defined in the air toxics sections of the Clean Air Act and thus emit less than 10 tons of a single toxic air pollutant or less than 25 tons of multiple toxic air pollutants in any one year. Area sources include smaller facilities, such as dry cleaners.

As a separate definition, area sources in air-quality modeling refer to those modeled in two dimensions (with length and width), as compared to point sources modeled at a single location.

Assessment System for Population Exposure Nationwide (ASPEN):

A computer simulation model used to estimate toxic air pollutant concentrations. The ASPEN model takes into account important determinants of pollutant concentrations, such as: rate of release, location of release, the height from which the pollutants are released, wind speeds and directions from the meteorological stations nearest to release, breakdown of the pollutants in the atmosphere after being released (i.e., reactive decay), settling of pollutants out of the atmosphere (i.e., deposition), and transformation of one pollutant into another (i.e., secondary formation or decay). The model estimates toxic air pollutant concentrations for every census tract in the United States, Puerto Rico, and the Virgin Islands.

Atmospheric transformation (secondary formation):

The process by which chemicals are transformed in the air into other chemicals. When a chemical is transformed, the original HAP no longer exists; it is replaced by one or more chemicals. Compared to the original chemical, the newer reaction products can have more, less, or the same toxicity. Transformations and removal processes affect both the fate of the chemical and its atmospheric persistence. Persistence is important because human exposure to chemical is influenced by the length of time the chemical remains in the atmosphere. Note that in NATA the terms atmospheric transformation and secondary formation are used interchangeably.

Background concentrations:

For NATA, the contributions to outdoor air toxics concentrations resulting from natural sources, persistence in the environment of past years' emissions, and long-range transport from distant sources. Background concentrations could be levels of pollutants that would be found in a particular year, even if there had been no recent manmade emissions. Background concentrations are added to the AERMOD concentrations but not to the CMAQ modeled concentrations which account for long range transport and emissions from outside the domain through boundary conditions. The vast majority of risk from the NATA background concentrations is from carbon tetrachloride, a ubiquitous pollutant that has few sources of emissions but is persistent due to its long half-life.

Biogenic emissions:

Emissions from natural sources, such as plants and trees. These sources emit formaldehyde, acetaldehyde, and methanol, as well as large quantities of other non-HAP volatile organic compounds (VOCs). Formaldehyde and acetaldehyde are key risk drivers in NATA. Biogenic emissions are typically computed using a model which utilizes spatial information on vegetation and land use and environmental conditions of temperature and solar radiation. In addition to being a primary source of HAPs, other VOCs emitted by biogenic sources react with anthropogenic VOCs and NO_X to produce secondary-formed HAPs. The NATA biogenics source group includes only the primary emissions.

Cancer risk:

The probability of contracting cancer over the course of a lifetime, assuming continuous exposure (assumed to be 70 years for the purposes of NATA risk characterization).

Carcinogen:

A chemical or physical agent that can cause cancer.

Chemical Abstracts Service (CAS) Number:

A unique number assigned to a chemical by the Chemical Abstracts Service, a service of the American Chemical Society that indexes and compiles abstracts of worldwide chemical literature called “Chemical Abstracts.” The purpose is to make database searches more convenient, as chemicals often have many names.

Census tracts:

Land areas defined by the U.S. Census Bureau. Tracts can vary in size but each typically contains about 4,000 residents. Census tracts are usually smaller than 2 square miles in cities, but are much larger in rural areas.

Cohort:

Generally defined as a group of people within a population who are assumed to have identical exposures during a specified exposure period. The use of cohorts is a necessary simplifying assumption for modeling exposures of a large population. For the exposure assessment, the population is divided into a set of cohorts such that (1) each person is assigned to one and only one cohort, and (2) all the cohorts combined encompass the entire population.

Community Multi-scale Air Quality (CMAQ) modeling system:

A multi-pollutant air quality modeling system using a three-dimensional gridded simulation environment with atmospheric chemistry to model transport of emissions across local to long-range scales.

Consolidated Human Activity Database (CHAD):

The Consolidated Human Activity Database (CHAD) is an EPA comprehensive human-activity database consisting of data from numerous activity studies since 1982 and supporting assessments of human exposure, intake dose, and risk.

Diesel particulate matter (diesel PM):

A mixture of particles that is a component of diesel exhaust. EPA lists diesel exhaust as a mobile-source air toxic due to the cancer and non-cancer health effects associated with exposure to whole diesel exhaust. Diesel PM (expressed as grams diesel PM/m^3) has historically been used as a surrogate measure of exposure for whole diesel exhaust. Although uncertainty exists as to whether diesel PM is the most appropriate parameter to correlate with human health effects, it is considered a reasonable choice until more definitive information about the mechanisms of toxicity or mode(s) of action of diesel exhaust becomes available.

Dispersion model:

A computerized set of mathematical equations that uses emissions and meteorological information to simulate the behavior and movement of air pollutants in the atmosphere. The results of a dispersion model are estimated outdoor concentrations of individual air pollutants at specified locations.

Emission Inventory System (EIS):

An EPA information system for storing all current and historical emission inventory data. It is used to receive and store emissions data and generate emission inventories beginning with the 2008 National Emissions Inventory (NEI). Partners used the EIS Exchange to submit Facility Inventory, Point, Nonpoint, Onroad and Nonroad data categories to the EIS Production or Quality Assurance (QA) environments.

Exposure assessment:

Identifying the ways in which chemicals might reach individuals (e.g., by breathing); estimating how much of a chemical an individual is likely to be exposed to; and, estimating the number of individuals likely to be exposed.

Hazard index (HI):

The sum of hazard quotients for substances that affect the same target organ or organ system. Because different pollutants (air toxics) can cause similar adverse health effects, combining hazard quotients associated with different substances is often appropriate. EPA has drafted revisions to the national guidelines on mixtures that support combining the effects of different substances in specific and limited ways. Ideally, hazard quotients should be combined for pollutants that cause adverse effects by the same toxic mechanism. Because detailed information on toxic mechanisms is not available for most of the substances in NATA, however, EPA aggregates the effects when they affect the same target organ regardless of the mechanism. The hazard index (HI) is only an approximation of the aggregate effect on the target organ (e.g., the lungs) because some of the substances might cause irritation by different (i.e., non-additive) mechanisms. As with the hazard quotient, aggregate exposures below an HI of 1.0 derived using target organ specific hazard quotients likely will not result in adverse non-cancer health effects over a lifetime of exposure and would ordinarily be considered acceptable. An HI equal to or greater than 1.0, however, does not necessarily suggest a likelihood of adverse effects. Because of the inherent conservatism of the reference concentration (RfC) methodology, the acceptability of exceedances must be evaluated on a case-by-case basis, considering such factors as the confidence level of the assessment, the size of the uncertainty factors used, the slope of the dose-response curve, the magnitude of the exceedance, and the number or types of people exposed at various levels above the RfC. Furthermore, the HI cannot be translated to a probability that adverse effects will occur, and it is not likely to be proportional to risk.

Hazard quotient (HQ):

The ratio of the potential exposure to the substance and the level at which no adverse effects are expected. A hazard quotient less than or equal to one indicates that adverse noncancer effects are not likely to occur, and thus can be considered to have negligible hazard. HQs greater than one are not statistical probabilities of harm occurring. Instead, they are a simple statement of whether (and by how much) an exposure concentration exceeds the reference concentration (RfC). Moreover, the level of concern does not increase linearly or to the same extent as HQs increase above one for different chemicals because RfCs do not generally have equal accuracy or precision and are generally not based on the same severity of effect. Thus, we can only say that with exposures increasingly greater than the RfC, (i.e., HQs increasingly greater than 1), the potential for adverse effects increases, but we do not know by how much. An HQ of 100 does not mean that the hazard is 10 times greater than an HQ of 10. Also an HQ of 10 for one substance may not have the same meaning (in terms of hazard) as another substance resulting in the same HQ.

Hazardous Air Pollutant Exposure Model (HAPEM):

A computer model that has been designed to estimate inhalation exposure for specified population groups and air toxics. Through a series of calculation routines, the model makes use of census data, human-activity patterns, ambient air quality levels, and indoor/outdoor concentration relationships to estimate an expected range of inhalation exposure concentrations for groups of individuals.

Human Exposure Model (HEM):

The Human Exposure Model (HEM) is a computer model used primarily for conducting inhalation risk assessments for sources emitting air toxics to ambient air. HEM-3 contains the AERMOD dispersion model for air-transport simulations and U.S. Census data for identifying population receptors.

Inhalation:

Breathing. Once inhaled, contaminants can be deposited in the lungs, taken into the blood, or both.

Integrated Risk Information System (IRIS):

The Integrated Risk Information System (IRIS) is an EPA program that identifies and characterizes the health hazards of chemicals found in the environment. IRIS is EPA's preferred source of toxicity information.

Lifetime cancer risk:

The probability of contracting cancer over the course of a lifetime (assumed to be 70 years for the purposes of NATA risk characterization).

Major sources:

Defined by the Clean Air Act as those stationary facilities that emit or have the potential to emit 10 tons of any one toxic air pollutant or 25 tons of more than one toxic air pollutant per year.

Maximum-likelihood estimate:

The most accurate maximum likelihood estimate is, by definition, the mode of a data set (i.e., the most frequent observation). When data are too limited to identify a clear mode, the average or the median of the data is usually substituted. For some air toxics for which adequate human data exist, EPA has based the unit risk estimate on the maximum-likelihood estimate for response data or for fitted curves.

Median:

The middle value of a set of ordered values (i.e., half the numbers are less than or equal to the median value). A median is the 50th percentile of the data.

Motor Vehicle Emission Simulator (MOVES):

A state-of-the-science emissions modeling system that estimates emissions for mobile sources at the national, county, and project level for criteria air pollutants, air toxics, and greenhouse gases.

Microenvironment:

A small space in which human contact with a pollutant takes place. A microenvironment can be treated as a well-characterized, relatively homogenous location with respect to pollutant concentrations for a specified period. For NATA, the Hazardous Air Pollutant Exposure Model considers cohort activities in 18 microenvironment locations that include (1) indoor locations (e.g., residence, office, store, school, restaurant, church, manufacturing facility, auditorium, healthcare facility, service station, other public building, garage); (2) outdoor locations (e.g., parking lot/garage, near road, motorcycle, service station, construction site, residential grounds, school, sports arena, park/golf course); and (3) in-vehicle locations (e.g., car, bus, truck, other, train/subway, airplane).

Microgram:

One-millionth of a gram. One gram is about one twenty-eighth of an ounce.

National-scale Air Toxics Assessment (NATA):

EPA's ongoing comprehensive evaluation of air toxics in the United States. These activities include the expansion of air toxics monitoring, improvement and periodic updating of emission inventories, improvement of national- and local-scale modeling, continued research on health effects and exposures to both ambient and indoor air, and improvement of assessment tools.

National Emissions Inventory (NEI):

EPA prepares a national database of air emissions information with input from numerous state and local air agencies, from tribes, and from industry. This database contains information on stationary and mobile sources that emit criteria air pollutants and their precursors, as well as hazardous air pollutants. The database includes estimates of annual emissions, by source, of air pollutants in each area of the country, on an annual basis. The National Emissions Inventory includes emission estimates for all 50 states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands.

National Mobile Inventory Model (NMIM):

Computer application containing EPA’s NONROAD model for estimating county level inventories of nonroad mobile emissions.

Noncancer risk:

The risk associated with effects other than cancer, based on the reference concentration, which is an estimate, with uncertainty spanning perhaps an order of magnitude, of an inhalation exposure to the human population (including sensitive subgroups) that is likely to be without appreciable risks of deleterious effects during a lifetime.

Nonroad mobile sources:

Mobile sources not found on roads and highways (e.g., airplanes, trains, lawn mowers, construction vehicles, farm machinery).

On-road mobile sources:

Vehicles found on roads and highways (e.g., cars, trucks, buses).

Percentile:

Any one of the points dividing a distribution of values into parts that each contain 1/100 of the values. For example, the 75th percentile is a value such that 75 percent of the values are less than or equal to it. In this assessment, the distribution of values represented (national, state, or county percentiles) depends on the presentation format of the results (map, bar chart, or data table).

Polycyclic organic matter (POM):

Defines a broad class of compounds that includes polycyclic aromatic hydrocarbons. Polycyclic organic matter (POM) compounds are formed primarily from combustion and are present in the atmosphere in particulate form. Sources of air emissions are diverse and include vehicle exhausts, forest fires and wildfires, asphalt roads, coal, coal tar, coke ovens, agricultural burning, residential wood burning, and hazardous waste sites. Not all POM reported to EPA's National Emission Inventory is speciated. As a result, EPA applies some simplifying assumptions to model and assess the risk from the individual pollutants that comprise polycyclic organic matter.

Reference concentration (RfC):

The reference concentration is an estimate (with uncertainty spanning perhaps an order of magnitude) of a continuous inhalation exposure to the human population (including sensitive subgroups that include children, asthmatics, and the elderly) that is likely to be without an appreciable risk of deleterious effects during a lifetime. It can be derived from various types of human or animal data, with uncertainty factors generally applied to reflect limitations of the data used.

Risk:

The probability that damage to life, health, or the environment will occur as a result of a given hazard (such as exposure to a toxic chemical). Some risks can be measured or estimated in numerical terms (e.g., one chance in a hundred).

Rural:

Consistent with the definition EPA used in the analyses to support the Integrated Urban Air Toxics Strategy, a county is considered “rural” if it does not contain a metropolitan statistical area with a population greater than 250,000 and the U.S. Census Bureau does not designate more than 50 percent of the population as “urban.” Note that this definition does not necessarily apply for any regulatory or implementation purpose.

Sparse Matrix Operator Kernel Emissions (SMOKE):

A modeling system that processes emissions data for use in gridded air quality models. It uses the Biogenic Emission Inventory System (BEIS) to model biogenic emissions. It also has a feature to use MOVES emission factors, activity data and meteorological data to compute hourly gridded onroad mobile emissions.

Science Advisory Board (SAB):

A panel of scientists, engineers, and economists who provide EPA with independent scientific and technical advice.

Stationary sources:

Emission sources other than mobile sources such as large industrial sources such as power plants and refineries, smaller industrial and commercial sources such as dry cleaners and commercial cooking, and residential sources such as residential wood combustion and consumer products usage. Stationary sources may be characterized as being emitted from “major“ sources or “area“ sources based on the 10-ton or 25-ton definitions contained in the Clean Air Act. For presentation purposes, the NATA results are identified as “point“ and “nonpoint“ sources rather than “major“ and “area“ sources. The point and nonpoint designations reflect the way each source of emissions is modeled. Some smaller sources that are area sources in the inventory (based on the amount of their emissions) are modeled as point sources because the location of their emissions was identified with latitude and longitude coordinates.

Susceptibility:

An increased likelihood of an adverse effect, often discussed in terms of relationship to a factor (e.g., life stage, demographic feature, or genetic characteristic) that can be used to describe a human subpopulation.

Toxicity weighting:

A relative risk evaluation tool that normalizes the emissions rates of each pollutant to a hypothetical substance with an inhalation unit risk value of 1/µg/m^3 (for carcinogenic effects) or a reference concentration of 1 mg/m^3 (for non-cancer effects). It is entirely emissions-based and toxicity-based, and does not consider dispersion, fate, receptor locations, and other exposure parameters. It may be calculated based on the emissions data for all pollutants released from a facility or source being assessed. It is particularly useful if the number of pollutants is large and the desire is to focus the risk analysis on a smaller subset of pollutants that contribute the most to risk.

Typical:

Describes a hypothetical person living at the census-tract centroid (defined as a reference point that is usually but not always located at the geographic center of a census tract) and engaging in a range of activities (indoors and outdoors) that are representative of those in which individuals residing in that tract might engage. To characterize the risk that this person might experience, NATA divides the population as a whole into cohorts (groups who are assumed to have identical exposures during a specified exposure period) based on where they live, how old they are, and what their daily-activity patterns might be. For each combination of residential census tract, age, various age-appropriate daily-activity patterns are selected to represent the range of exposure conditions for residents of the tract. A population-weighted typical exposure estimate is calculated for each cohort, and this value is used to estimate representative risks for a "typical" individual residing in that tract.

Upper bound:

A plausible upper limit to the true value of a quantity; usually not a true statistical confidence limit.

Upper-bound lifetime cancer risk:

A plausible upper limit to the true probability that an individual will contract cancer over a 70-year lifetime as a result of a given hazard (such as exposure to a toxic chemical). This risk can be measured or estimated in numerical terms (e.g., one chance in a hundred).

Unit risk estimate (URE):

The upper-bound excess lifetime cancer risk estimated to result from continuous exposure to an agent at a concentration of 1 µg/m^3 in air. The interpretation of the unit risk estimate (URE) would be as follows: If the URE = 1.5×10^{-6} per µg/m^3, 1.5 excess tumors are expected to develop per 1,000,000 people if they were exposed daily for a lifetime to 1 µg of the chemical in 1 m^3 of air. UREs are considered upper-bound estimates, meaning they represent a plausible upper limit to the true value. (Note that this is usually not a true statistical confidence limit.) The true risk is likely to be less, but could be greater.

Urban:

Consistent with the definition EPA used in the analyses to support the Integrated Urban Air Toxics Strategy, a county is considered "urban" if it either includes a metropolitan statistical area with a population greater than 250,000 or the U.S. Census Bureau designates more than 50 percent of the population as "urban." Note that this definition does not necessarily apply for any regulatory or implementation purpose.

Weight-of-evidence (WOE) for carcinogenicity:

The weight-of-evidence (WOE) narrative for carcinogenicity is a summary that explains what is known about an agent's human carcinogenic potential and the conditions that characterize its expression. The narrative should be sufficiently complete to stand alone, highlighting the key issues and decisions that were the basis for the evaluation of the agent's potential hazard. The WOE characterizes the extent to which the available data support the hypothesis that an agent causes cancer in humans. Under EPA's 1986 risk assessment guidelines, the weight of evidence is described by categories "A through E," with Group A for known human carcinogens through Group E for agents with evidence of non-carcinogenicity. The approach outlined in EPA's guidelines for carcinogen risk assessment (2005) considers all scientific information in determining if and under what conditions an agent can cause cancer in humans, and provides a narrative approach to characterize carcinogenicity rather than categories. To provide clarity and consistency in an otherwise free-form, narrative characterization, standard descriptors are used as part of the hazard narrative to express the conclusion regarding the WOE for carcinogenic hazard potential. Five standard hazard descriptors are recommended: (1) carcinogenic to humans, (2) likely to be carcinogenic to humans, (3) suggestive evidence of carcinogenic potential, (4) inadequate information to assess carcinogenic potential, and (5) not likely to be carcinogenic to humans.

Carcinogenic to humans: This descriptor indicates strong evidence of human carcinogenicity. It covers different combinations of evidence. This descriptor is appropriate when the epidemiologic evidence of a causal association between human exposure and cancer is convincing. An exception is that this descriptor might also be equally appropriate with a lesser weight of epidemiologic evidence that is strengthened by other lines of evidence. This descriptor can be used when all of the following conditions are met: (a) there is strong evidence of an association between human exposure and either cancer or the key precursor events of the agent's mode of action but not enough for a causal association; (b) there is extensive evidence of carcinogenicity in animals; (c) the mode(s) of carcinogenic action and associated key precursor events have been identified in animals, (d) there is strong evidence that the key precursor events that precede the cancer response in animals are anticipated to occur in humans and progress to tumors, based on available biological information.

Likely to be carcinogenic to humans: This descriptor is appropriate when the weight of the evidence is adequate to demonstrate carcinogenic potential to humans but does not reach the WOE for the descriptor "carcinogenic to humans." Adequate evidence consistent with this descriptor covers a broad spectrum. At one end of the spectrum is evidence for an association between human exposure to the agent and cancer and strong experimental evidence of carcinogenicity in animals; at the other, with no human data, the weight of experimental evidence shows animal carcinogenicity by a mode or modes of action that are relevant or assumed to be relevant to humans. The use of the term "likely" as a WOE descriptor does not correspond to a quantifiable probability. Moreover, additional information, for example, on mode of action, might change the choice of descriptor for the illustrated examples.

Suggestive evidence of carcinogenic potential: This descriptor is appropriate when the WOE suggests carcinogenicity; a concern for potential carcinogenic effects in humans is raised, but the data are judged insufficient for a stronger conclusion. This descriptor covers a spectrum of evidence associated with varying levels of concern for carcinogenicity, ranging from a positive cancer result in the only study on an agent to a single positive cancer result in an extensive data base that includes negative studies in other species. Depending on the extent of the data base, additional studies might or might not provide further insights.

Inadequate information to assess carcinogenic potential: This descriptor is appropriate when available data are judged inadequate for applying one of the other descriptors. Additional studies generally would be expected to provide further insights.

Not likely to be carcinogenic to humans: This descriptor is appropriate when the available data are considered robust for deciding that there is no basis for human hazard concern. In some instances, there can be positive results in experimental animals when the evidence is strong and consistent that each mode of action in experimental animals does not operate in humans. In other cases, the evidence in both humans and animals that the agent is not carcinogenic can be convincing. "Not likely" applies only to the circumstances supported by the data. For example, an agent might be "not likely to be carcinogenic" by one route but not necessarily by another. In cases having positive animal experiment(s) but the results are judged not to be relevant to humans, the narrative discusses why the results are not relevant.

Weather Research and Forecasting (WRF) model:

A mesoscale numerical weather-prediction system for atmospheric research and weather forecasting. It can generate atmospheric conditions using real input data or idealized conditions.

Appendix B

Air Toxics Included in Modeling for the 2011 NATA, and Source Classification Codes that Define Diesel Particulate Matter

This appendix contains three tables. The first two are related to the air toxics included in the 2011 NATA, and the third lists the source classification codes (SCC) for which the PM_{10} emissions were considered to be diesel particulate matter (PM).

Exhibit B-1 contains the air toxics included in the 2011 NATA and indicates the inventory types(s) reporting them. The names shown in this table match the terminology used in the 1990 Clean Air Act (CAA) Amendments; for example, this table lists “chromium compounds“ but does not indicate which individual compounds containing chromium were modeled, and it lists four forms of xylenes (o-, m-, p- and mixed isomers) but these were grouped and modeled as a single entity. See Appendix C for the names of the actual substances included in the 2011 NATA. Exhibit B-1 also contains indications about whether cancer risks and chronic non-cancer hazard quotients were estimated for each air toxic. Appendix H provides the toxicity values used in NATA.

Exhibit B-2 contains the air toxics that were *not* modeled for the 2011 NATA and why. Note that although diesel PM was modeled for NATA and is included in Exhibit B-1, it is not categorized as a HAP in the CAA. Diesel PM emissions were computed based on PM_{10} emissions from onroad and nonroad mobile sources burning diesel or residual fuels (see Exhibit B-3).

The excel file “NATA_Pollutants_AppendixB_AppendixC.xlsx” in the SupplementalData folder provides the data in spreadsheet format and includes additional fields such as the CMAQ model species names.

Note that NEI = National Emissions Inventory.

Exhibit B-1. Air Toxics Included in NATA

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Data Category in the NEI												Background [j]	Assessed for Cancer	Assessed for Noncancer
		CMAQ	Secondary	Event [b]	Nonpoint [c]	Nonroad	Onroad	Point [d]	Ag burning [e]	Rail yards [f]	Airports [g]	Locomotives [h]	CMV [i]			
1,1,2,2-Tetrachloroethane	79345	✓			✓			✓						NC		
1,1,2-Trichloroethane	79005				✓			✓						Y	✓	✓
1,1-Dimethyl hydrazine	57147							✓								
1,2,4-Trichlorobenzene	120821				✓			✓								✓
1,2-Dibromo-3-chloropropane	96128							✓						Y	✓	✓
1,2-Diphenylhydrazine [k]	122667							✓							✓	
1,2-Epoxybutane	106887				✓			✓								✓
1,2-Propylenimine (2-methyl aziridine)	75558							✓								
1,3-Butadiene	106990	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	NC	✓	✓
1,3-Dichloropropene	542756	✓			✓			✓						NC	✓	✓
1,3-Propane sultone	1120714							✓							✓	
1,4-Dichlorobenzene(p)	106467	✓			✓			✓						NC	✓	✓
1,4-Dioxane	123911				✓			✓							✓	✓
2,2,4-Trimethylpentane	540841				✓	✓	✓	✓		✓	✓	✓	✓	Y		
2,4,5-Trichlorophenol	95954							✓								
2,4,6-Trichlorophenol	88062				✓			✓							✓	
2,4-D, salts and esters	94757				✓			✓								
2,4-Dinitrophenol	51285				✓			✓								
2,4-Dinitrotoluene	121142				✓			✓							✓	✓
2,4-Toluene diamine	95807							✓							✓	
2,4-Toluene diisocyanate	584849	✓			✓			✓							✓	✓
2- Acetylaminofluorene	53963							✓							✓	
2-Chloroacetophenone	532274				✓			✓								✓
2-Nitropropane	79469				✓			✓							✓	✓
3,3'-Dichlorobenzidine	91941							✓							✓	
3,3'-Dimethoxybenzidine	119904							✓								
3,3'-Dimethylbenzidine	119937							✓								
4,4'-Methylene bis(2-chloroaniline)	101144							✓							✓	
4,4'-Methylenedianiline	101779							✓							✓	✓
4,6-Dinitro-o-cresol, and salts	534521							✓								
4-Aminobiphenyl	92671							✓								
4-Nitrobiphenyl	92933							✓								
4-Nitrophenol	100027				✓			✓								
Acetaldehyde	75070	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓
Acetamide	60355				✓			✓							✓	
Acetonitrile	75058				✓			✓						Y		✓
Acetophenone	98862				✓			✓								
Acrolein	107028	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓
Acrylamide	79061				✓			✓							✓	✓
Acrylic acid	79107				✓			✓								✓
Acrylonitrile	107131	✓			✓			✓						NC	✓	✓

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Data Category in the NEI												Background [j]	Assessed for Cancer	Assessed for Noncancer
		CMAQ	Secondary	Event [b]	Nonpoint [c]	Nonroad	Onroad	Point [d]	Ag burning [e]	Rail yards [f]	Airports [g]	Locomotives [h]	CMV [i]			
Allyl chloride	107051				✓			✓							✓	✓
Aniline	62533							✓							✓	✓
Antimony Compounds	7440360				✓			✓				✓	✓	Y		✓
Arsenic Compounds (inorganic including arsine)	7440382	✓			✓	✓	✓	✓	✓	✓		✓	✓	NC	✓	✓
Benzene (including benzene from gasoline)	71432	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	NC	✓	✓
Benzidine	92875							✓						Y	✓	✓
Benzotrichloride	98077							✓								
Benzyl chloride	100447				✓			✓						Y	✓	
Beryllium Compounds	7440417	✓			✓			✓		✓		✓	✓	NC	✓	✓
Beta-Propiolactone [k]	57578							✓								
Biphenyl	92524				✓			✓								
Bis(2-ethylhexyl)phthalate (DEHP)	117817				✓			✓						Y	✓	✓
Bis(chloromethyl)ether	542881							✓							✓	
Bromoform	75252				✓			✓						Y	✓	
Cadmium Compounds	7440439	✓			✓			✓	✓	✓		✓	✓	NC	✓	✓
Calcium cyanamide	156627							✓								✓
Captan	133062				✓			✓								
Carbaryl	63252				✓			✓								
Carbon disulfide	75150				✓			✓						Y		✓
Carbon tetrachloride	56235	✓			✓			✓						Y	✓	✓
Carbonyl sulfide	463581			✓	✓			✓								✓
Catechol	120809							✓								
Chloramben [k]	133904							✓								
Chlordane	57749							✓							✓	✓
Chlorine	7782505	✓			✓			✓	✓			✓	✓			✓
Chloroacetic acid	79118							✓								
Chlorobenzene	108907				✓			✓								✓
Chlorobenzilate	510156							✓							✓	
Chloroform	67663	✓			✓			✓						NC		✓
Chloromethyl methyl ether	107302							✓								
Chloroprene	126998				✓			✓							✓	✓
Chromium Compounds [l]	multiple	✓			✓	✓	✓	✓						NC		
Cobalt Compounds	7440484				✓			✓		✓		✓	✓	Y		✓
Coke Oven Emissions	140							✓							✓	
Cresols/Cresylic acid (isomers and mixture) [m]	1319773				✓			✓								✓
Cumene	98828				✓			✓		✓	✓	✓	✓	Y		✓
Cyanide Compounds	multiple				✓			✓								✓
Diazomethane	334883															
Dibenzofurans	132649				✓			✓								
Dibutylphthalate	84742				✓			✓								

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Data Category in the NEI												Background [j]	Assessed for Cancer	Assessed for Noncancer
		CMAQ	Secondary	Event [b]	Nonpoint [c]	Nonroad	Onroad	Point [d]	Ag burning [e]	Rail yards [f]	Airports [g]	Locomotives [h]	CMV [i]			
Dichloroethyl ether (Bis(2-chloroethyl)ether)	111444							✓							✓	
Dichlorvos	62737							✓								✓
Diethanolamine	111422				✓			✓								✓
Diethyl sulfate	64675							✓								
Dimethyl aminoazobenzene	60117							✓							✓	
Dimethyl carbamoyl chloride	79447							✓								
Dimethyl formamide	68122				✓			✓								✓
Dimethyl phthalate	131113				✓			✓								
Dimethyl sulfate	77781				✓			✓								
Epichlorohydrin (l-Chloro-2,3-epoxypropane)	106898				✓			✓							✓	✓
Ethyl acrylate	140885				✓			✓								
Ethyl benzene	100414				✓	✓	✓	✓		✓	✓	✓	✓	Y	✓	✓
Ethyl carbamate (Urethane)	51796							✓							✓	
Ethyl chloride (Chloroethane)	75003				✓			✓								✓
Ethylene dibromide (Dibromoethane)	106934	✓			✓			✓		✓				NC	✓	✓
Ethylene dichloride (1,2-Dichloroethane)	107062	✓			✓			✓		✓				NC	✓	✓
Ethylene glycol	107211				✓			✓								✓
Ethylene imine (Aziridine)	151564							✓								
Ethylene oxide	75218	✓			✓			✓						NC	✓	✓
Ethylene thiourea	96457							✓							✓	✓
Ethylidene dichloride (1,1-Dichloroethane)	75343				✓			✓							✓	✓
Formaldehyde	50000	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓
Glycol Ethers	N/A				✓			✓		✓						✓
Heptachlor	76448							✓							✓	
Hexachlorobenzene	118741				✓			✓					✓		✓	✓
Hexachlorobutadiene	87683				✓			✓							✓	✓
Hexachlorocyclopentadiene	77474				✓			✓								✓
Hexachloroethane	67721							✓								✓
Hexamethylene-1,6-diisocyanate	822060	✓			✓			✓								✓
Hexamethylphosphoramide [k]	680319															
Hexane	110543			✓	✓	✓	✓	✓		✓	✓	✓	✓	Y		✓
Hydrazine	302012	✓						✓							✓	✓
Hydrochloric acid	7647010	✓			✓			✓								✓
Hydrogen fluoride (Hydrofluoric acid)	7664393				✓			✓								✓
Hydroquinone	123319				✓			✓								
Isophorone	78591				✓			✓								✓
Lead Compounds	7439921	✓			✓			✓	✓	✓	✓	✓	✓	NC		✓
Lindane (all isomers)	58899				✓			✓							✓	✓

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Data Category in the NEI												Background [j]	Assessed for Cancer	Assessed for Noncancer
		CMAQ	Secondary	Event [b]	Nonpoint [c]	Nonroad	Onroad	Point [d]	Ag burning [e]	Rail yards [f]	Airports [g]	Locomotives [h]	CMV [i]			
Maleic anhydride	108316	✓			✓			✓								✓
Manganese Compounds	7439965	✓			✓	✓	✓	✓	✓	✓		✓	✓	NC		✓
m-Cresol [m]	108394							✓								✓
Mercury Compounds	7439976	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓	NC		✓
Methanol	67561	✓			✓			✓	✓	✓	✓	✓	✓			✓
Methoxychlor	72435							✓								
Methyl bromide (Bromomethane)	74839				✓			✓						Y		✓
Methyl chloride (Chloromethane)	74873			✓	✓			✓						Y		✓
Methyl chloroform (1,1,1-Trichloroethane)	71556				✓			✓						Y		✓
Methyl hydrazine	60344				✓			✓								
Methyl iodide (Iodomethane)	74884				✓			✓								
Methyl isobutyl ketone (Hexone)	108101				✓			✓		✓				Y		✓
Methyl isocyanate	624839							✓								✓
Methyl methacrylate	80626				✓			✓								✓
Methyl tert butyl ether	1634044				✓			✓							✓	✓
Methylene chloride (Dichloromethane)	75092	✓			✓			✓						NC	✓	✓
Methylene diphenyl diisocyanate (MDI)	101688				✓			✓								✓
m-Xylenes [n]	108383	✓		✓	✓	✓	✓	✓		✓	✓	✓	✓	NC		✓
N,N-Dimethylaniline	121697				✓			✓								
Naphthalene	91203	✓			✓	✓	✓	✓		✓	✓	✓	✓	NC	✓	✓
Nickel Compounds	7440020	✓			✓	✓	✓	✓	✓	✓		✓	✓	NC	✓	✓
Nitrobenzene	98953				✓			✓							✓	✓
N-Nitrosodimethylamine	62759							✓							✓	
N-Nitrosomorpholine	59892							✓							✓	
N-Nitroso-N-Methylurea [k]	684935							✓								
o-Anisidine	90040							✓								
o-Cresol [m]	95487				✓			✓								✓
o-Toluidine	95534				✓			✓							✓	
o-Xylenes [n]	95476	✓		✓	✓	✓	✓	✓		✓	✓	✓	✓	NC		✓
Parathion [k]	56382							✓								
p-Cresol [m]	106445				✓			✓								✓
Pentachloronitrobenzene (Quintobenzene)	82688				✓			✓								
Pentachlorophenol	87865				✓			✓							✓	✓
Phenol	108952				✓			✓			✓					✓
Phosgene	75445							✓								✓
Phosphine	7803512							✓								✓
Phosphorus	7723140				✓			✓				✓	✓			
Phthalic anhydride	85449				✓			✓								✓

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Data Category in the NEI												Background [j]	Assessed for Cancer	Assessed for Noncancer
		CMAQ	Secondary	Event [b]	Nonpoint [c]	Nonroad	Onroad	Point [d]	Ag burning [e]	Rail yards [f]	Airports [g]	Locomotives [h]	CMV [i]			
Polychlorinated biphenyls (Aroclors)	1336363				✓			✓					✓		✓	
Polycyclic Organic Matter [o]	N/A	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	
p-Phenylenediamine	106503							✓								
Propionaldehyde	123386				✓	✓	✓	✓		✓	✓	✓	✓	Y		✓
Propoxur (Baygon)	114261							✓								
Propylene dichloride (1,2-Dichloropropane)	78875	✓			✓			✓						NC		✓
Propylene oxide	75569				✓			✓							✓	✓
p-Xylenes [n]	106423	✓		✓	✓	✓	✓	✓		✓		✓	✓	NC		✓
Quinoline	91225	✓						✓								
Quinone	106514							✓								
Selenium Compounds	7782492				✓			✓		✓		✓	✓	Y		✓
Styrene	100425				✓	✓	✓	✓		✓	✓	✓	✓	Y		✓
Styrene oxide	96093				✓			✓								✓
Tetrachloroethylene (Perchloroethylene)	127184	✓			✓			✓						NC	✓	✓
Titanium tetrachloride	7550450							✓								✓
Toluene	108883	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	NC		✓
Toxaphene (chlorinated camphene)	8001352							✓							✓	
Trichloroethylene	79016	✓			✓			✓						NC	✓	✓
Triethylamine	121448	✓			✓			✓								✓
Trifluralin	1582098				✓			✓								
Vinyl acetate	108054				✓			✓		✓				Y		✓
Vinyl bromide	593602							✓							✓	✓
Vinyl chloride	75014	✓			✓			✓		✓				NC	✓	✓
Vinylidene chloride (1,1-Dichloroethylene)	75354				✓			✓								✓
Xylenes (isomers and mixture) [n]	1330207	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	NC		✓
Diesel PM [p]		✓								✓	✓	✓	✓			✓

[a] In most cases, the NEI pollutant code is the same as the CAS number. In a few cases (e.g., coke oven emissions) a CAS number has not been assigned, and NEI uses a unique pollutant code.

Note: Actual categories in the NEI are: point, nonpoint, onroad, nonroad, and event. The above categories break out point and nonpoint into additional groups. See footnotes for more details

[b] Event category has prescribed and wildfires (day-specific)

[c] Excluding ag fires, locomotives, and commercial marine vessels

[d] Excluding airports and rail yards

[e] Ag burning is agricultural field burning and is a part of the nonpoint data category in the NEI

[f] Railyards are part of the point data category in the NEI

[g] Airports are part of the point data category in the NEI

[h] CMV = commercial marine vessels, part of the nonpoint category in the NEI

[i] Locomotives are part of the nonpoint category in the NEI

[j] Remote concentration estimate added: NC="non-CONUS" (concentration added only to non-continental U.S. areas: Alaska, Hawaii, Puerto Rico, and U.S. Virgin Islands), Y=non-CMAQ air toxic remote concentration estimate added everywhere

[k] Not in 2011 NATA because there were no emissions

[l] NATA includes only hexavalent chromium

[m] Modeled as cresols

[n] Modeled as xylenes

[o] About 50 specific compounds are in the NEI. They were modeled as 9 discrete PAH groups representing different URE "bins" since specific compounds have a wide range of UREs.

[p] Diesel PM is not a HAP and not on the Clean Air Act list but it is modeled in NATA.

Exhibit B-2. Pollutants Excluded from NATA

Pollutant	NEI Pollutant Code (CAS Number) [a]	Reason for Exclusion	In Previous NATAs?
2,3,7,8-Tetrachlorodibenzo-p-dioxin	1746016	Dioxins and furans are not in the 2011 NEI due to uncertainty in the completeness or accuracy of the S/L/T agency data for this group of pollutants. In addition, the most significant exposure route for dioxin is ingestion, not inhalation, so dioxin's relative contribution to NATA's inhalation risk estimates likely would not be large.	n
Other dioxins/furans	multiple		n
Radionuclides		Radionuclides are not in the 2011 NEI due to uncertainty in the completeness or accuracy of the S/L/T agency data for this group of pollutants. In addition, the NEI currently is not compatible with emissions reported in units other than mass, and therefore suitable emissions data have not been compiled for these substances on a national scale.	n
DDE	72559 incorrectly referred to in the Section 112(b) list as 3547-04-4	This pollutant was not reported to the 2011 NEI.	y
Fine mineral fibers (including rockwool and slag wool and fine mineral fibers)	Fine mineral fibers: 383 Rockwool:617 Slagwool:616	Rockwool has 0 emissions and slagwool and fine mineral fibers are excluded from previous assessments	n
Asbestos	1332214	Inhalation exposures not typically expressed in mass units	n
Diazomethane	334883	This pollutant has 0 emissions in the 2011 NEI	y
Hexamethylphosphoramide	680319	This pollutant has 0 emissions in the 2011 NEI	n

[a] In most cases, the NEI pollutant code is the same as the CAS number. In a few cases (e.g., coke oven emissions) a CAS number has not been assigned, and NEI uses a unique pollutant code.

Exhibit B-3. Source Classification Codes For which PM_{10} Emissions were assigned to Diesel Particulate Matter

NEI Category	SCC	Description
Point	28500201	Internal Combustion Engines;Railroad Equipment;Diesel;Yard Locomotives
	2270008005	*M;Off-highway Vehicle Diesel;Airport Ground Support Equipment;Airport Ground Support Equipment
Nonpoint	2280002100	*M;Marine Vessels, Commercial;Diesel;Port emissions
	2280002200	*M;Marine Vessels, Commercial;Diesel;Underway emissions
	2285002006	*M;Railroad Equipment;Diesel;Line Haul Locomotives: Class I Operations
	2285002007	*M;Railroad Equipment;Diesel;Line Haul Locomotives: Class II / III Operations
	2285002008	*M;Railroad Equipment;Diesel;Line Haul Locomotives: Passenger Trains (Amtrak)
	2285002009	*M;Railroad Equipment;Diesel;Line Haul Locomotives: Commuter Lines
	2285002010	*M;Railroad Equipment;Diesel;Yard Locomotives
	2280003100	*M;Marine Vessels, Commercial;Residual;Port emissions
	2280003200	*M;Marine Vessels, Commercial;Residual;Underway emissions
Nonroad	2270001060	*M;Off-highway Vehicle Diesel;Recreational Equipment;Specialty Vehicles/Carts
	2270002003	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Pavers
	2270002006	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Tampers/Rammers
	2270002009	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Plate Compactors
	2270002015	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Rollers
	2270002018	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Scrapers
	2270002021	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Paving Equipment
	2270002024	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Surfacing Equipment
	2270002027	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Signal Boards/Light Plants
	2270002030	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Trenchers
	2270002033	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Bore/Drill Rigs
	2270002036	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Excavators
	2270002039	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Concrete/Industrial Saws
	2270002042	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Cement and Mortar Mixers
	2270002045	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Cranes
	2270002048	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Graders
	2270002051	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Off-highway Trucks
	2270002054	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Crushing/Processing Equipment
	2270002057	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Rough Terrain Forklifts
	2270002060	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Rubber Tire Loaders

NEI Category	SCC	Description
	2270002066	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Tractors/Loaders/Backhoes
	2270002069	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Crawler Tractor/Dozers
	2270002072	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Skid Steer Loaders
	2270002075	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Off-highway Tractors
	2270002078	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Dumpers/Tenders
	2270002081	*M;Off-highway Vehicle Diesel;Construction and Mining Equipment;Other Construction Equipment
	2270003010	*M;Off-highway Vehicle Diesel;Industrial Equipment;Aerial Lifts
	2270003020	*M;Off-highway Vehicle Diesel;Industrial Equipment;Forklifts
	2270003030	*M;Off-highway Vehicle Diesel;Industrial Equipment;Sweepers/Scrubbers
	2270003040	*M;Off-highway Vehicle Diesel;Industrial Equipment;Other General Industrial Equipment
	2270003050	*M;Off-highway Vehicle Diesel;Industrial Equipment;Other Material Handling Equipment
	2270003060	*M;Off-highway Vehicle Diesel;Industrial Equipment;ACRefrigeration
	2270003070	*M;Off-highway Vehicle Diesel;Industrial Equipment;Terminal Tractors
	2270004031	*M;Off-highway Vehicle Diesel;Lawn and Garden Equipment;Leafblowers/Vacuums (Commercial)
	2270004036	*M;Off-highway Vehicle Diesel;Lawn and Garden Equipment;Snowblowers (Commercial)
	2270004046	*M;Off-highway Vehicle Diesel;Lawn and Garden Equipment;Front Mowers (Commercial)
	2270004056	*M;Off-highway Vehicle Diesel;Lawn and Garden Equipment;Lawn and Garden Tractors (Commercial)
	2270004066	*M;Off-highway Vehicle Diesel;Lawn and Garden Equipment;Chippers/Stump Grinders (Commercial)
	2270004071	*M;Off-highway Vehicle Diesel;Lawn and Garden Equipment;Turf Equipment (Commercial)
	2270004076	*M;Off-highway Vehicle Diesel;Lawn and Garden Equipment;Other Lawn and Garden Equipment (Commercial)
	2270005010	*M;Off-highway Vehicle Diesel;Agricultural Equipment;2-Wheel Tractors
	2270005015	*M;Off-highway Vehicle Diesel;Agricultural Equipment;Agricultural Tractors
	2270005020	*M;Off-highway Vehicle Diesel;Agricultural Equipment;Combines
	2270005025	*M;Off-highway Vehicle Diesel;Agricultural Equipment;Balers
	2270005030	*M;Off-highway Vehicle Diesel;Agricultural Equipment;Agricultural Mowers
	2270005035	*M;Off-highway Vehicle Diesel;Agricultural Equipment;Sprayers
	2270005040	*M;Off-highway Vehicle Diesel;Agricultural Equipment;Tillers > 6 HP
	2270005045	*M;Off-highway Vehicle Diesel;Agricultural Equipment;Swathers
	2270005055	*M;Off-highway Vehicle Diesel;Agricultural Equipment;Other Agricultural Equipment
	2270005060	*M;Off-highway Vehicle Diesel;Agricultural Equipment;Irrigation Sets
	2270006005	*M;Off-highway Vehicle Diesel;Commercial Equipment;Generator Sets
	2270006010	*M;Off-highway Vehicle Diesel;Commercial Equipment;Pumps
	2270006015	*M;Off-highway Vehicle Diesel;Commercial Equipment;Air Compressors

NEI Category	SCC	Description
	2270006020	*M;Off-highway Vehicle Diesel;Commercial Equipment;Gas Compressors
	2270006025	*M;Off-highway Vehicle Diesel;Commercial Equipment;Welders
	2270006030	*M;Off-highway Vehicle Diesel;Commercial Equipment;Pressure Washers
	2270006035	*M;Off-highway Vehicle Diesel;Commercial Equipment;Hydro-power Units! new SCC in 2002v2
	2270007010	*M;Off-highway Vehicle Diesel;Logging Equipment;Shredders > 6 HP
	2270007015	*M;Off-highway Vehicle Diesel;Logging Equipment;Forest Eqp - Feller/Bunch/Skidder
	2270009010	*M;Off-highway Vehicle Diesel;Underground Mining Equipment;Other Underground Mining Equipment
	2270010010	*M;Off-highway Vehicle Diesel;Industrial Equipment;Other Oil Field Equipment
	2282020005	*M;Pleasure Craft;Diesel;Inboard/Sterndrive
	2282020010	*M;Pleasure Craft;Diesel;Outboard
	2285002015	*M;Railroad Equipment;Diesel;Railway Maintenance
Onroad	2202210181	*D;Passenger Cars;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202210281	*D;Passenger Cars;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202210381	*D;Passenger Cars;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202210481	*D;Passenger Cars;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202210581	*D;Passenger Cars;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202310181	*D;Passenger Trucks;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202310281	*D;Passenger Trucks;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202310381	*D;Passenger Trucks;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202310481	*D;Passenger Trucks;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202310581	*D;Passenger Trucks;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202320181	*D;Light Commercial Trucks;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202320281	*D;Light Commercial Trucks;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202320381	*D;Light Commercial Trucks;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202320481	*D;Light Commercial Trucks;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202320581	*D;Light Commercial Trucks;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202410181	*D;Intercity Buses;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202410281	*D;Intercity Buses;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202410381	*D;Intercity Buses;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202410481	*D;Intercity Buses;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202410581	*D;Intercity Buses;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202420181	*D;Transit Buses;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202420281	*D;Transit Buses;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling

NEI Category	SCC	Description
	2202420381	*D;Transit Buses;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202420481	*D;Transit Buses;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202420581	*D;Transit Buses;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202430181	*D;School Buses;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202430281	*D;School Buses;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202430381	*D;School Buses;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202430481	*D;School Buses;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202430581	*D;School Buses;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202510181	*D;Refuse Trucks;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202510281	*D;Refuse Trucks;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202510381	*D;Refuse Trucks;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202510481	*D;Refuse Trucks;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202510581	*D;Refuse Trucks;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202520181	*D;Single Unit Short-haul Trucks;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202520281	*D;Single Unit Short-haul Trucks;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202520381	*D;Single Unit Short-haul Trucks;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202520481	*D;Single Unit Short-haul Trucks;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202520581	*D;Single Unit Short-haul Trucks;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202530181	*D;Single Unit Long-haul Trucks;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202530281	*D;Single Unit Long-haul Trucks;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202530381	*D;Single Unit Long-haul Trucks;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202530481	*D;Single Unit Long-haul Trucks;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202530581	*D;Single Unit Long-haul Trucks;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202540181	*D;Motor Homes;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202540281	*D;Motor Homes;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202540381	*D;Motor Homes;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202540481	*D;Motor Homes;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202540581	*D;Motor Homes;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202610181	*D;Combination Short-haul Trucks;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202610281	*D;Combination Short-haul Trucks;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling

NEI Category	SCC	Description
	2202610381	*D;Combination Short-haul Trucks;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202610481	*D;Combination Short-haul Trucks;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202610581	*D;Combination Short-haul Trucks;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202620153	*D;Combination Long-haul Trucks;Off-network: Extended Idle Exhaust
	2202620181	*D;Combination Long-haul Trucks;Off-network: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202620191	*D;Combination Long-haul Trucks;Off-network: Auxiliary Power Exhaust
	2202620281	*D;Combination Long-haul Trucks;Rural Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202620381	*D;Combination Long-haul Trucks;Rural Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202620481	*D;Combination Long-haul Trucks;Urban Restricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling
	2202620581	*D;Combination Long-haul Trucks;Urban Unrestricted Access: All Exhaust, Evaporative, Brake, and Tire Except Refueling and Hoteling

*M= Mobile Sources. *D=Highway Vehicles–Diesel

Appendix C

Crosswalk for Air Toxics Names in the NEI and Metal Speciation Factors

Exhibit C-1 contains the air toxic name crosswalk and metal speciation factors used to conduct the modeling of emissions for the 2011 NATA. This crosswalk contains a link between lists of air toxic names in two data bases used for NATA:

- the names used in the National Emissions Inventory (NEI), and
- the names used for NATA.

This table also contains the corresponding names for each air toxic as used in the 1990 Clean Air Act Amendments. In addition, Exhibit C-1 contains the speciation of metal chemicals based on their metal mass fractions.

The metal speciation factor was used to adjust modeled mass emissions prior to modeling and conducting risk calculations, because metal toxicity is usually evaluated relative to the amount of metal ion present rather than the total mass of the metal compound. Most metal and cyanide compounds are reported in the 2011 NEI as just the metal or cyanide parts; consequently, most fractions are 1, including the two cyanide compounds. If the NEI data reporters did not adjust the emissions downward to account for just the metal part, a more health-protective (higher risk) result would be obtained.

A master pollutant list for NATA in spreadsheet format, "NATA_Pollutants_AppendixB_AppendixC.xlsx", is provided in the SupplementalData folder. The second sheet in the workbook is an electronic form of the PAHPOM groupings shown in Section 2.1.1.2 of the TSD.

Exhibit C-1. Crosswalk for Air Toxics Names in NEI, NATA Results, and the Clean Air Act, with Metal Speciation Factors

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Glycol Ethers	112072	Glycol Ethers	2-Butoxyethyl Acetate	GLYCOL ETHERS	1
Glycol Ethers	112152	Glycol Ethers	Carbitol Acetate	GLYCOL ETHERS	1
Glycol Ethers	112254	Glycol Ethers	2-(Hexyloxy)Ethanol	GLYCOL ETHERS	1
Glycol Ethers	112276	Glycol Ethers	Triethylene glycol	GLYCOL ETHERS	1
Glycol Ethers	112345	Glycol Ethers	Diethylene Glycol Monobutyl Ether	GLYCOL ETHERS	1
Glycol Ethers	112356	Glycol Ethers	Methoxytriglycol	GLYCOL ETHERS	1
Glycol Ethers	112367	Glycol Ethers	Diethylene Glycol Diethyl Ether	GLYCOL ETHERS	1
Glycol Ethers	112492	Glycol Ethers	Triethylene Glycol Dimethyl Ether	GLYCOL ETHERS	1
Glycol Ethers	112594	Glycol Ethers	N-Hexyl Carbitol	GLYCOL ETHERS	1
Propoxur (Baygon)	114261	Propoxur	Propoxur	PROPOXUR (BAYGON)	1
Bis(2-ethylhexyl)phthalate (DEHP)	117817	Bis(2-Ethylhexyl)Phthalate	Bis(2-Ethylhexyl)Phthalate	BIS(2-ETHYLHEXYL)PHTHALATE (DEHP)	1
Hexachlorobenzene	118741	Hexachlorobenzene	Hexachlorobenzene	HEXACHLOROBENZENE	1
3,3'-Dimethoxybenzidine	119904	3,3'-Dimethoxybenzidine	3,3'-Dimethoxybenzidine	3,3'-DIMETHOXYBENZIDINE	1
3,3'-Dimethylbenzidine	119937	3,3'-Dimethylbenzidine	3,3'-Dimethylbenzidine	3,3'-DIMETHYLYBENZIDINE	1
Polycyclic Organic Matter	120127	Polycyclic Organic Matter	Anthracene	PAHPOM	1
Catechol	120809	Catechol	Catechol	CATECHOL	1
1,2,4-Trichlorobenzene	120821	1,2,4-Trichlorobenzene	1,2,4-Trichlorobenzene	1,2,4-TRICHLOROBENZENE	1
2,4-Dinitrotoluene	121142	2,4-Dinitrotoluene	2,4-Dinitrotoluene	2,4-DINITROTOLUENE	1
Triethylamine	121448	Triethylamine	Triethylamine	TRIETHYLAMINE	1
N,N-Dimethylaniline	121697	N,N-Dimethylaniline	N,N-Dimethylaniline	N,N-DIMETHYLANILINE	1
1,2-Diphenylhydrazine	122667	1,2-Diphenylhydrazine	1,2-Diphenylhydrazine	1,2-DIPHENYLHYDRAZINE	1
Glycol Ethers	122996	Glycol Ethers	Phenyl Cellosolve	GLYCOL ETHERS	1
4-Nitrophenol	100027	4-Nitrophenol	4-Nitrophenol	4-NITROPHENOL	1
Ethyl benzene	100414	Ethylbenzene	Ethyl Benzene	ETHYLBENZENE	1
Styrene	100425	Styrene	Styrene	STYRENE	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Benzyl chloride	100447	Benzyl Chloride	Benzyl Chloride	BENZYL CHLORIDE	1
4,4'-Methylene bis(2-chloroaniline)	101144	4,4'-Methylenebis(2-Chloroaniline)	4,4'-Methylenebis(2-Chloraniline)	4,4'-METHYLENE BIS(2-CHLOROANILINE)	1
Methylene diphenyl diisocyanate (MDI)	101688	4,4'-Methylenediphenyl Diisocyanate	4,4'-Methylenediphenyl Diisocyanate	4,4'-METHYLENEDIPHENYL DIISOCYANATE (MDI)	1
4,4'-Methylenedianiline	101779	4,4'-Methylenedianiline	4,4'-Methylenedianiline	4,4'-METHYLENEDIANILINE	1
p-Xylenes	106423	Xylenes (Mixed Isomers)	p-Xylene	XYLENES (MIXED ISOMERS)	1
p-Cresol	106445	Cresol/Cresylic Acid (Mixed Isomers)	p-Cresol	CRESOL_CRESYLIC ACID (MIXED ISOMERS)	1
1,4-Dichlorobenzene(p)	106467	1,4-Dichlorobenzene	1,4-Dichlorobenzene	1,4-DICHLOROBENZENE	1
p-Phenylenediamine	106503	p-Phenylenediamine	p-Phenylenediamine	P-PHENYLENEDIAMINE	1
Quinone	106514	Quinone	Quinone	QUINONE (P-BENZOQUINONE)	1
1,2-Epoxybutane	106887	1,2-Epoxybutane	1,2-Epoxybutane	1,2-EPOXYBUTANE	1
Epichlorohydrin (l-Chloro-2,3-epoxypropane)	106898	Epichlorohydrin	Epichlorohydrin	EPICHLOROHYDRIN	1
Ethylene dibromide (Dibromoethane)	106934	Ethylene Dibromide	Ethylene Dibromide	ETHYLENE DIBROMIDE (DIBROMOETHANE)	1
1,3-Butadiene	106990	1,3-Butadiene	1,3-Butadiene	1,3-BUTADIENE	1
Acrolein	107028	Acrolein	Acrolein	ACROLEIN	1
Allyl chloride	107051	Allyl Chloride	Allyl Chloride	ALLYL CHLORIDE	1
Ethylene dichloride (1,2-Dichloroethane)	107062	Ethylene Dichloride	Ethylene Dichloride	ETHYLENE DICHLORIDE (1,2-DICHLOROETHANE)	1
Acrylonitrile	107131	Acrylonitrile	Acrylonitrile	ACRYLONITRILE	1
Ethylene glycol	107211	Ethylene Glycol	Ethylene Glycol	ETHYLENE GLYCOL	1
Chloromethyl methyl ether	107302	Chloromethyl Methyl Ether	Chloromethyl Methyl Ether	CHLOROMETHYL METHYL ETHER	1
Vinyl acetate	108054	Vinyl Acetate	Vinyl Acetate	VINYL ACETATE	1
Methyl isobutyl ketone (Hexone)	108101	Methyl Isobutyl Ketone	Methyl Isobutyl Ketone	METHYL ISOBUTYL KETONE (HEXONE)	1
Maleic anhydride	108316	Maleic Anhydride	Maleic Anhydride	MALEIC ANHYDRIDE	1
m-Xylenes	108383	Xylenes (Mixed Isomers)	m-Xylene	XYLENES (MIXED ISOMERS)	1
m-Cresol	108394	Cresol/Cresylic Acid (Mixed Isomers)	m-Cresol	CRESOL_CRESYLIC ACID (MIXED ISOMERS)	1
Toluene	108883	Toluene	Toluene	TOLUENE	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Chlorobenzene	108907	Chlorobenzene	Chlorobenzene	CHLOROBENZENE	1
Phenol	108952	Phenol	Phenol	PHENOL	1
Glycol Ethers	109864	Glycol Ethers	Ethylene Glycol Methyl Ether	GLYCOL ETHERS	1
Glycol Ethers	110496	Glycol Ethers	Ethylene Glycol Monomethyl Ether Acetate	GLYCOL ETHERS	1
Hexane	110543	Hexane	Hexane	HEXANE	1
Glycol Ethers	110714	Glycol Ethers	1,2-Dimethoxyethane	GLYCOL ETHERS	1
Glycol Ethers	110805	Glycol Ethers	Cellosolve Solvent	GLYCOL ETHERS	1
Glycol Ethers	111159	Glycol Ethers	Cellosolve Acetate	GLYCOL ETHERS	1
Diethanolamine	111422	Diethanolamine	Diethanolamine	DIETHANOLAMINE	1
Dichloroethyl ether (Bis(2-chloroethyl)ether)	111444	Dichloroethyl Ether	Dichloroethyl Ether	DICHLOROETHYL ETHER (BIS[2-CHLOROETHYL]ETHER)	1
Glycol Ethers	111773	Glycol Ethers	Diethylene Glycol Monomethyl Ether	GLYCOL ETHERS	1
Glycol Ethers	111900	Glycol Ethers	Diethylene Glycol Monoethyl Ether	GLYCOL ETHERS	1
Glycol Ethers	111966	Glycol Ethers	Diethylene Glycol Dimethyl Ether	GLYCOL ETHERS	1
1,3-Propane sultone	1120714	1,3-Propane Sultone	1,3-Propanesultone	1,3-PROPANE SULTONE	1
Hydroquinone	123319	Hydroquinone	Hydroquinone	HYDROQUINONE	1
Propionaldehyde	123386	Propionaldehyde	Propionaldehyde	PROPIONALDEHYDE	1
1,4-Dioxane	123911	p-Dioxane	p-Dioxane	1,4-DIOXANE	1
Glycol Ethers	124174	Glycol Ethers	Butyl Carbitol Acetate	GLYCOL ETHERS	1
Chloroprene	126998	Chloroprene	Chloroprene	CHLOROPRENE	1
Tetrachloroethylene (Perchloroethylene)	127184	Tetrachloroethylene	Tetrachloroethylene	TETRACHLOROETHYLENE	1
Polycyclic Organic Matter	129000	Polycyclic Organic Matter	Pyrene	PAHPOM	1
Polycyclic Organic Matter	85018	Polycyclic Organic Matter	Phenanthrene	PAHPOM	1
Dimethyl phthalate	131113	Dimethyl Phthalate	Dimethyl Phthalate	DIMETHYL PHTHALATE	1
Nickel Compounds	1313991	Nickel Compounds	Nickel Oxide	NICKEL COMPOUNDS	0.7412
Cresols/Cresylic acid (isomers and mixture)	1319773	Cresol/Cresylic Acid (Mixed Isomers)	Cresol/Cresylic Acid (Mixed Isomers)	CRESOL_CRESYLIC ACID (MIXED ISOMERS)	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Dibenzofurans	132649	Dibenzofuran	Dibenzofuran	DIBENZOFURAN	1
Xylenes (isomers and mixture)	1330207	Xylenes (Mixed Isomers)	Xylenes (Mixed Isomers)	XYLENES (MIXED ISOMERS)	1
Captan	133062	Captan	Captan	CAPTAN	1
Asbestos	1332214	Asbestos	Asbestos	Not used in NATA	1
Chromium Compounds	1333820	Chromium Compounds	Chromium Trioxide	CHROMIUM VI (HEXAVALENT)	0.52
Polychlorinated biphenyls (Aroclors)	1336363	Polychlorinated Biphenyls	Polychlorinated Biphenyls	POLYCHLORINATED BIPHENYLS (AROCLORS)	1
Chloramben	133904	Chloramben	Chloramben	Chloramben	1
Coke Oven Emissions	140	Coke Oven Emissions	Coke Oven Emissions	COKE OVEN EMISSIONS	1
Ethyl acrylate	140885	Ethyl Acrylate	Ethyl Acrylate	ETHYL ACRYLATE	1
Glycol Ethers	143226	Glycol Ethers	Triglycol Monobutyl Ether	GLYCOL ETHERS	1
Ethylene imine (Aziridine)	151564	Ethyleneimine (Aziridine)	Ethyleneimine	ETHYLENEIMINE (AZIRIDINE)	1
Calcium cyanamide	156627	Calcium Cyanamide	Calcium Cyanamide	CALCIUM CYANAMIDE	0.5
Trifluralin	1582098	Trifluralin	Trifluralin	TRIFLURALIN	1
Chromium Compounds	16065831	Chromium Compounds	Chromium III	Not used in NATA	1
Methyl tert butyl ether	1634044	Methyl Tert-Butyl Ether	Methyl Tert-Butyl Ether	METHYL TERT-BUTYL ETHER	1
Glycol Ethers	16672392	Glycol Ethers	Di(Ethylene Glycol Monobutyl Ether) Phthalate	GLYCOL ETHERS	1
Glycol Ethers	171	Glycol Ethers	Glycol Ethers	GLYCOL ETHERS	1
Chromium Compounds	18540299	Chromium Compounds	Chromium (VI)	CHROMIUM VI (HEXAVALENT)	1
Polycyclic Organic Matter	86748	Polycyclic Organic Matter	Carbazole	PAHPOM	1
Polycyclic Organic Matter	218019	Polycyclic Organic Matter	Chrysene	PAHPOM	1
Polycyclic Organic Matter	130498292	Polycyclic Organic Matter	PAH, total	PAHPOM	1
Polycyclic Organic Matter	191242	Polycyclic Organic Matter	Benzo[g,h,i,]Perylene	PAHPOM	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Polycyclic Organic Matter	192972	Polycyclic Organic Matter	Benzo[e]Pyrene	PAHPOM	1
Polycyclic Organic Matter	195197	Polycyclic Organic Matter	Benzo(c)phenanthrene	PAHPOM	1
Polycyclic Organic Matter	198550	Polycyclic Organic Matter	Perylene	PAHPOM	1
Polycyclic Organic Matter	206440	Polycyclic Organic Matter	Fluoranthene	PAHPOM	1
Polycyclic Organic Matter	208968	Polycyclic Organic Matter	Acenaphthylene	PAHPOM	1
Polycyclic Organic Matter	2381217	Polycyclic Organic Matter	1-Methylpyrene	PAHPOM	1
Polycyclic Organic Matter	2422799	Polycyclic Organic Matter	12-Methylbenz(a)Anthracene	PAHPOM	1
Polycyclic Organic Matter	250	Polycyclic Organic Matter	PAH/POM - Unspecified	PAHPOM	1
Polychlorinated Biphenyls	2050682	Polychlorinated Biphenyls	4,4'-Dichlorobiphenyl (PCB-15)	POLYCHLORINATED BIPHENYLS (AROCLORS)	1
Polychlorinated Biphenyls	2051243	Polychlorinated Biphenyls	Decachlorobiphenyl (PCB-209)	POLYCHLORINATED BIPHENYLS (AROCLORS)	1
Polychlorinated Biphenyls	2051607	Polychlorinated Biphenyls	2-Chlorobiphenyl (PCB-1)	POLYCHLORINATED BIPHENYLS (AROCLORS)	1
Polycyclic Organic Matter	26914181	Polycyclic Organic Matter	Methylanthracene	PAHPOM	1
Polycyclic Organic Matter	65357699	Polycyclic Organic Matter	Methylbenzopyrene	PAHPOM	1
Polycyclic Organic Matter	8007452	Polycyclic Organic Matter	Coal Tar	PAHPOM	1
Glycol Ethers	20706256	Glycol Ethers	2-Propoxyethyl Acetate	GLYCOL ETHERS	1
Polycyclic Organic Matter	832699	Polycyclic Organic Matter	1-Methylphenanthrene	PAHPOM	1
Polycyclic Organic Matter	83329	Polycyclic Organic Matter	Acenaphthene	PAHPOM	1
Polycyclic Organic Matter	86737	Polycyclic Organic Matter	Fluorene	PAHPOM	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Polycyclic Organic Matter	90120	Polycyclic Organic Matter	1-Methylnaphthalene	PAHPOM	1
Polycyclic Organic Matter	91576	Polycyclic Organic Matter	2-Methylnaphthalene	PAHPOM	1
Polycyclic Organic Matter	91587	Polycyclic Organic Matter	2-Chloronaphthalene	PAHPOM	1
Polycyclic Organic Matter	193395	Polycyclic Organic Matter	Indeno[1,2,3-c,d]Pyrene	PAHPOM	1
Polycyclic Organic Matter	203123	Polycyclic Organic Matter	Benzo(g,h,i)Fluoranthene	PAHPOM	1
Polychlorinated Biphenyls	25429292	Polychlorinated Biphenyls	Pentachlorobiphenyl	POLYCHLORINATED BIPHENYLS (AROCLORS)	1
Polychlorinated Biphenyls	26601649	Polychlorinated Biphenyls	Hexachlorobiphenyl	POLYCHLORINATED BIPHENYLS (AROCLORS)	1
Polycyclic Organic Matter	203338	Polycyclic Organic Matter	Benzo(a)Fluoranthene	PAHPOM	1
Polychlorinated Biphenyls	26914330	Polychlorinated Biphenyls	Tetrachlorobiphenyl	POLYCHLORINATED BIPHENYLS (AROCLORS)	1
Glycol Ethers	2807309	Glycol Ethers	Propyl Cellosolve	GLYCOL ETHERS	1
Polycyclic Organic Matter	284	POM as non-15 PAH	Extractable Organic Matter (EOM)	PAHPOM	1
Polychlorinated Biphenyls	28655712	Polychlorinated Biphenyls	Heptachlorobiphenyl	POLYCHLORINATED BIPHENYLS (AROCLORS)	1
Hydrazine	302012	Hydrazine	Hydrazine	HYDRAZINE	1
Diazomethane	334883	Diazomethane	Diazomethane	No emissions in 2011, so not in 2011 NATA)	1
Polycyclic Organic Matter	205823	Polycyclic Organic Matter	Benzo[j]fluoranthene	PAHPOM	1
Fine Mineral Fibers	383	Fine Mineral Fibers	Fine Mineral Fibers	Not used in NATA	1
Polycyclic Organic Matter	205992	Polycyclic Organic Matter	Benzo[b]Fluoranthene	PAHPOM	1
Carbonyl sulfide	463581	Carbonyl Sulfide	Carbonyl Sulfide	CARBONYL SULFIDE	1
Formaldehyde	50000	Formaldehyde	Formaldehyde	FORMALDEHYDE	1
Polycyclic Organic Matter	207089	Polycyclic Organic Matter	Benzo[k]Fluoranthene	PAHPOM	1
Chlorobenzilate	510156	Chlorobenzilate	Chlorobenzilate	CHLOROBENZILATE	1
2,4-Dinitrophenol	51285	2,4-Dinitrophenol	2,4-Dinitrophenol	2,4-DINITROPHENOL	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Ethyl carbamate (Urethane)	51796	Ethyl Carbamate	Ethyl Carbamate	ETHYL CARBAMATE (URETHANE) CHLORIDE (CHLOROETHANE)	1
2-Chloroacetophenone	532274	2-Chloroacetophenone	2-Chloroacetophenone	2-CHLOROACETOPHENONE	1
4,6-Dinitro-o-cresol, and salts	534521	4,6-Dinitro-o-Cresol	4,6-Dinitro-o-Cresol	4,6-DINITRO-O-CRESOL (INCLUDING SALTS)	1
Polycyclic Organic Matter	224420	Polycyclic Organic Matter	Dibenzo[a,j]Acridine	PAHPOM	1
2-Acetylaminofluorene	53963	2-Acetylaminofluorene	2-Acetylaminofluorene	2-ACETYLAMINOFLUORENE	1
2,2,4-Trimethylpentane	540841	2,2,4-Trimethylpentane	2,2,4-Trimethylpentane	2,2,4-TRIMETHYLPENTANE	1
1,3-Dichloropropene	542756	1,3-Dichloropropene	1,3-Dichloropropene	1,3-DICHLOROPROPENE	1
Bis(chloromethyl)ether	542881	Bis(Chloromethyl) Ether	Bis(Chloromethyl)Ether	BIS(CHLOROMETHYL) ETHER	1
Polycyclic Organic Matter	226368	Polycyclic Organic Matter	Dibenz[a,h]acridine	PAHPOM	1
Carbon tetrachloride	56235	Carbon Tetrachloride	Carbon Tetrachloride	CARBON TETRACHLORIDE	1
Parathion	56382	Parathion	Parathion	Parathion	1
Polycyclic Organic Matter	5522430	Polycyclic Organic Matter	1-Nitropyrene	PAHPOM	1
Polycyclic Organic Matter	56553	Polycyclic Organic Matter	Benz[a]Anthracene	PAHPOM	1
Polycyclic Organic Matter	56832736	Polycyclic Organic Matter	Benzofluoranthenes	PAHPOM	1
Cyanide Compounds	57125	Cyanide Compounds	Cyanide	CYANIDE COMPOUNDS	1
1,1-Dimethyl hydrazine	57147	1,1-Dimethylhydrazine	1,1-Dimethyl Hydrazine	1,1-DIMETHYLHYDRAZINE	1
Beta-Propiolactone	57578	Beta-Propiolactone	Beta-Propiolactone	Beta-Propiolactone	1
Chlordane	57749	Chlordane	Chlordane	CHLORDANE	1
Polycyclic Organic Matter	192654	Polycyclic Organic Matter	Dibenzo[a,e]Pyrene	PAHPOM	1
2,4-Toluene diisocyanate	584849	2,4-Toluene Diisocyanate	2,4-Toluene Diisocyanate	2,4-TOLUENE DIISOCYANATE	1
Lindane (all isomers)	58899	Lindane (All isomers)	1,2,3,4,5,6-Hexachlorocyclohexane	1,2,3,4,5,6-HEXACHLOROCYCLYHEXANE	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Vinyl bromide	593602	Vinyl Bromide	Vinyl Bromide	VINYL BROMIDE	1
N-Nitrosomorpholine	59892	N-Nitrosomorpholine	N-Nitrosomorpholine	N-NITROSOMORPHOLINE	1
4-Dimethylaminoazobenzene	60117	4-Dimethylaminoazobenzene	4-Dimethylaminoazobenzene	4-DIMETHYLAMINOAZOBENZENE	1
Methyl hydrazine	60344	Methylhydrazine	Methylhydrazine	METHYLHYDRAZINE	1
Acetamide	60355	Acetamide	Acetamide	ACETAMIDE	1
Nickel Compounds	604	Nickel Compounds	Nickel Refinery Dust	NICKEL COMPOUNDS	1
Fine Mineral Fibers	616	Fine Mineral Fibers	Slagwool (Man-Made Fibers)	Not used in NATA	1
Fine Mineral Fibers	617	Fine Mineral Fibers	Rockwool (Man-Made Fibers)	Not used in NATA	1
Methyl isocyanate	624839	Methyl Isocyanate	Methyl Isocyanate	METHYL ISOCYANATE	1
Aniline	62533	Aniline	Aniline	ANILINE	1
Dichlorvos	62737	Dichlorvos	Dichlorvos	DICHLORVOS	1
N-Nitrosodimethylamine	62759	N-Nitrosodimethylamine	N-Nitrosodimethylamine	N-NITROSODIMETHYLAMINE	1
Carbaryl	63252	Carbaryl	Carbaryl	CARBARYL	1
Diethyl sulfate	64675	Diethyl Sulfate	Diethyl Sulfate	DIETHYL SULFATE	1
Polycyclic Organic Matter	194592	Polycyclic Organic Matter	7H-Dibenzo[c,g]carbazole	PAHPOM	1
Glycol Ethers	67425	Glycol Ethers	(Ethylenebis(Oxyethylenenitrilo)) Tetraacetic Acid	GLYCOL ETHERS	1
Methanol	67561	Methanol	Methanol	METHANOL	1
Chloroform	67663	Chloroform	Chloroform	CHLOROFORM	1
Hexachloroethane	67721	Hexachloroethane	Hexachloroethane	HEXACHLOROETHANE	1
Hexamethylphosphoramide	680319	Hexamethylphosphoramide	Hexamethylphosphoramide	no emissions in 2011, not in 2011 NATA	1
Dimethyl formamide	68122	N,N-Dimethylformamide	N,N-Dimethylformamide	DIMETHYL FORMAMIDE	1
N-Nitroso-N-Methylurea	684935	N-Nitroso-N-Methylurea	N-Nitroso-N-Methylurea	N-Nitroso-N-Methylurea	1
Polychlorinated Biphenyls	7012375	Polychlorinated Biphenyls	2,4,4'-Trichlorobiphenyl (PCB-28)	POLYCHLORINATED BIPHENYLS (AROCLORS)	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Benzene (including benzene from gasoline)	71432	Benzene	Benzene	BENZENE	1
Methyl chloroform (1,1,1-Trichloroethane)	71556	Methyl Chloroform	Methyl Chloroform	1,1,1-TRICHLOROETHANE	1
Methoxychlor	72435	Methoxychlor	Methoxychlor	METHOXYCHLOR	1
Manganese Compounds	7439965	Manganese Compounds	Manganese	MANGANESE COMPOUNDS	1
Mercury Compounds	7439976	Mercury Compounds	Mercury	MERCURY COMPOUNDS	1
Nickel Compounds	7440020	Nickel Compounds	Nickel	NICKEL COMPOUNDS	1
Antimony	7440360	Antimony Compounds	Antimony	ANTIMONY COMPOUNDS	1
Arsenic	7440382	Arsenic Compounds	Arsenic	ARSENIC COMPOUNDS(INORGANIC INCLUDING ARSINE)	1
Beryllium	7440417	Beryllium Compounds	Beryllium	BERYLLIUM COMPOUNDS	1
Cadmium	7440439	Cadmium Compounds	Cadmium	CADMIUM COMPOUNDS	1
Cobalt	7440484	Cobalt Compounds	Cobalt	COBALT COMPOUNDS	1
Methyl bromide (Bromomethane)	74839	Methyl Bromide	Methyl Bromide	METHYL BROMIDE (BROMOMETHANE)	1
Methyl chloride (Chloromethane)	74873	Methyl Chloride	Methyl Chloride	METHYL CHLORIDE (CHLOROMETHANE)	1
Methyl iodide (Iodomethane)	74884	Methyl Iodide	Methyl Iodide	METHYL IODIDE (IODOMETHANE)	1
Cyanide Compounds	74908	Cyanide Compounds	Hydrogen Cyanide	CYANIDE COMPOUNDS	1
Ethyl chloride (Chloroethane)	75003	Ethyl Chloride	Ethyl Chloride	ETHYL CHLORIDE	1
Vinyl chloride	75014	Vinyl Chloride	Vinyl Chloride	VINYL CHLORIDE	1
Acetonitrile	75058	Acetonitrile	Acetonitrile	ACETONITRILE	1
Acetaldehyde	75070	Acetaldehyde	Acetaldehyde	ACETALDEHYDE	1
Methylene chloride (Dichloromethane)	75092	Methylene Chloride	Methylene Chloride	METHYLENE CHLORIDE	1
Carbon disulfide	75150	Carbon Disulfide	Carbon Disulfide	CARBON DISULFIDE	1
Ethylene oxide	75218	Ethylene Oxide	Ethylene Oxide	ETHYLENE OXIDE	1
Bromoform	75252	Bromoform	Bromoform	BROMOFORM	1
Ethylidene dichloride (1,1-Dichloroethane)	75343	Ethylidene Dichloride	Ethylidene Dichloride	ETHYLIDENE DICHLORIDE (1,1-DICHLOROETHANE)	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Vinylidene chloride (1,1-Dichloroethylene)	75354	Vinylidene Chloride	Vinylidene Chloride	VINYLIDENE CHLORIDE	1
Phosgene	75445	Phosgene	Phosgene	PHOSGENE	1
Titanium tetrachloride	7550450	Titanium Tetrachloride	Titanium Tetrachloride	TITANIUM TETRACHLORIDE	1
1,2-Propylenimine (2-methyl aziridine)	75558	1,2-Propylenimine	1,2-Propylenimine	1,2-PROPYLENEIMINE	1
Propylene oxide	75569	Propylene Oxide	Propylene Oxide	PROPYLENE OXIDE	1
Heptachlor	76448	Heptachlor	Heptachlor	HEPTACHLOR	1
Hydrochloric acid	7647010	Hydrochloric Acid	Hydrochloric Acid	HYDROCHLORIC ACID (HYDROGEN CHLORIDE [GAS ONLY])	1
Hydrogen fluoride	7664393	Hydrogen Fluoride	Hydrogen Fluoride	HYDROGEN FLUORIDE (HYDROFLUORIC ACID)	1
Phosphorus	7723140	Phosphorus	Phosphorus	PHOSPHORUS	1
Chromium Compounds	7738945	Chromium Compounds	Chromic Acid (VI)	CHROMIUM VI (HEXAVALENT)	0.4406
Hexachlorocyclopentadiene	77474	Hexachlorocyclopentadiene	Hexachlorocyclopentadiene	HEXACHLOROCYCLOPENTADIENE	1
Dimethyl sulfate	77781	Dimethyl Sulfate	Dimethyl Sulfate	DIMETHYL SULFATE	1
Selenium Compounds	7782492	Selenium Compounds	Selenium	SELENIUM COMPOUNDS	1
Chlorine	7782505	Chlorine	Chlorine	CHLORINE	1
Glycol Ethers	7795917	Glycol Ethers	Ethylene Glycol Mono-Sec-Butyl Ether	GLYCOL ETHERS	1
Phosphine	7803512	Phosphine	Phosphine	PHOSPHINE	1
Isophorone	78591	Isophorone	Isophorone	ISOPHORONE	1
Propylene dichloride (1,2-Dichloropropane)	78875	Propylene Dichloride	Propylene Dichloride	PROPYLENE DICHLORIDE (1,2-DICHLOROPROPANE)	1
1,1,2-Trichloroethane	79005	1,1,2-Trichloroethane	1,1,2-Trichloroethane	1,1,2-TRICHLOROETHANE	1
Trichloroethylene	79016	Trichloroethylene	Trichloroethylene	TRICHLOROETHYLENE	1
Acrylamide	79061	Acrylamide	Acrylamide	ACRYLAMIDE	1
Acrylic acid	79107	Acrylic Acid	Acrylic Acid	ACRYLIC ACID	1
Chloroacetic acid	79118	Chloroacetic Acid	Chloroacetic Acid	CHLOROACETIC ACID	1
1,1,2,2-Tetrachloroethane	79345	1,1,2,2-Tetrachloroethane	1,1,2,2-Tetrachloroethane	1,1,2,2-TETRACHLOROETHANE	1
Dimethyl carbamoyl chloride	79447	Dimethylcarbamoyl Chloride	Dimethylcarbamoyl Chloride	DIMETHYLCARBAMOYL CHLORIDE	1
2-Nitropropane	79469	2-Nitropropane	2-Nitropropane	2-NITROPROPANE	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Toxaphene	8001352	Toxaphene	Toxaphene	TOXAPHENE (CHLORINATED CAMPHENE)	1
Polycyclic Organic Matter	3697243	Polycyclic Organic Matter	5-Methylchrysene	PAHPOM	1
Methyl methacrylate	80626	Methyl Methacrylate	Methyl Methacrylate	METHYL METHACRYLATE	1
Hexamethylene-1,6-diisocyanate	822060	Hexamethylene Diisocyanate	Hexamethylene Diisocyanate	HEXAMETHYLENE DIISOCYANATE	1
Pentachloronitrobenzene (Quintobenzene)	82688	Pentachloronitrobenzene	Pentachloronitrobenzene	PENTACHLORONITROBENZENE (QUINTOBENZENE)	1
Polycyclic Organic Matter	41637905	Polycyclic Organic Matter	Methylchrysene	PAHPOM	1
Polycyclic Organic Matter	50328	Polycyclic Organic Matter	Benzo[a]Pyrene	PAHPOM	1
Dibutylphthalate	84742	Dibutyl Phthalate	Dibutyl Phthalate	DIBUTYLPHTHALATE	1
Polycyclic Organic Matter	53703	Polycyclic Organic Matter	Dibenzo[a,h]Anthracene	PAHPOM	1
Phthalic anhydride	85449	Phthalic Anhydride	Phthalic Anhydride	PHTHALIC ANHYDRIDE	1
Polycyclic Organic Matter	56495	Polycyclic Organic Matter	3-Methylcholanthrene	PAHPOM	1
Polycyclic Organic Matter	189559	Polycyclic Organic Matter	Dibenzo[a,i]Pyrene	PAHPOM	1
Hexachlorobutadiene	87683	Hexachlorobutadiene	Hexachlorobutadiene	HEXACHLOROBUTADIENE	1
Pentachlorophenol	87865	Pentachlorophenol	Pentachlorophenol	PENTACHLOROPHENOL	1
2,4,6-Trichlorophenol	88062	2,4,6-Trichlorophenol	2,4,6-Trichlorophenol	2,4,6-TRICHLOROPHENOL	1
o-Anisidine	90040	o-Anisidine	o-Anisidine	ANISIDINE	1
Polycyclic Organic Matter	189640	Polycyclic Organic Matter	Dibenzo[a,h]Pyrene	PAHPOM	1
Naphthalene	91203	Naphthalene	Naphthalene	NAPHTHALENE	1
Quinoline	91225	Quinoline	Quinoline	QUINOLINE	1
Polycyclic Organic Matter	191300	Polycyclic Organic Matter	Dibenzo[a,l]Pyrene	PAHPOM	1
Polycyclic Organic Matter	57976	Polycyclic Organic Matter	7,12-Dimethylbenz[a]Anthracene	PAHPOM	1
3,3'-Dichlorobenzidine	91941	3,3'-Dichlorobenzidine	3,3'-Dichlorobenzidine	3,3'-DICHLOROBENZIDINE	1
Biphenyl	92524	Biphenyl	Biphenyl	BIPHENYL	1
4-Aminobiphenyl	92671	4-Aminobiphenyl	4-Aminobiphenyl	4-AMINOBIPHENYL	1

Air Toxic (Clean Air Act Name)	NEI Pollutant Code (CAS Number) [a]	Pollutant Category Name	Pollutant Description	NATA Website Pollutant Name	Metal Speciation Factor
Benzidine	92875	Benzidine	Benzidine	BENZIDINE	1
4-Nitrobiphenyl	92933	4-Nitrobiphenyl	4-Nitrobiphenyl	4-NITROBIPHENYL	1
2,4-D, salts and esters	94757	2,4-Dichlorophenoxy Acetic Acid	2,4-Dichlorophenoxy Acetic Acid	2,4-D, SALTS AND ESTERS	1
o-Xylenes	95476	Xylenes (Mixed Isomers)	o-Xylene	XYLENES (MIXED ISOMERS)	1
o-Cresol	95487	Cresol/Cresylic Acid (Mixed Isomers)	o-Cresol	CRESOL_CRESYLIC ACID (MIXED ISOMERS)	1
o-Toluidine	95534	o-Toluidine	o-Toluidine	O-TOLUIDINE	1
2,4-Toluene diamine	95807	Toluene-2,4-Diamine	Toluene-2,4-Diamine	2,4-TOLUENE DIAMINE	1
2,4,5-Trichlorophenol	95954	2,4,5-Trichlorophenol	2,4,5-Trichlorophenol	2,4,5-TRICHLOROPHENOL	1
Styrene oxide	96093	Styrene Oxide	Styrene Oxide	STYRENE OXIDE	1
1,2-Dibromo-3-chloropropane	96128	1,2-Dibromo-3-Chloropropane	1,2-Dibromo-3-Chloropropane	1,2-DIBROMO-3-CHLOROPROPANE	1
Ethylene thiourea	96457	Ethylene Thiourea	Ethylene Thiourea	ETHYLENE THIOUREA	1
Benzotrichloride	98077	Benzotrichloride	Benzotrichloride	BENZOTRICHLORIDE	1
Cumene	98828	Cumene	Cumene	CUMENE	1
Acetophenone	98862	Acetophenone	Acetophenone	ACETOPHENONE	1
Nitrobenzene	98953	Nitrobenzene	Nitrobenzene	NITROBENZENE	1
Lead Compounds	7439921	Lead Compounds	Lead	LEAD COMPOUNDS	1
Polycyclic Organic Matter	779022	Polycyclic Organic Matter	9-Methyl Anthracene	PAHPOM	1
Diesel PM [b]			Diesel PM	DIESEL PM	1

[a] In most cases, the NEI pollutant code is the same as the CAS number. In a few cases (e.g., coke oven emissions) a CAS number has not been assigned, and NEI uses a unique pollutant code.

[b] Diesel PM is not a Clean Air Act HAP.

This page intentionally left blank.

Appendix D

Additional Information Used to Process the 2011 NATA Inventory: Inventory Sectors and Model Run Groups; SCC Groupings; Speciations for Mercury, Xylenes, and Other Metals

This appendix contains additional information related to the emissions processing and categorization of the results. As described in Section 2 of the TSD, emission sources were grouped into Human Exposure Model (HEM) run groups and then disaggregated (using source classification code [SCC] information) into source groups. Both source groups and HEM run groups were categories used to group the results for the 2011 NATA.

Exhibit D-1 contains indications of how the source groups related to National Emissions Inventory (NEI) and Emission Inventory System (EIS) sectors.

Exhibit D-2 contains the SCCs for each source group. A spreadsheet file ("NP_NR_OR_SourceGroup_to_SCC_CrossReference.xlsx") containing a cross-reference file with SCC descriptions is provided in the SupplementalData folder. The point-source groups (point, airports, rail yards) were created based on the facility source type (100 for airports, 151 for rail yards).

Mercury and other metals and xylenes from the inventory were speciated for use in modeling. Mercury in the inventory was reported as pollutant code 7439976 and needs to be speciated into the three forms for CMAQ: elemental, divalent gaseous, and divalent particulate. Xylenes (mixed isomers) were speciated into m-, o- and p-forms, and metals (other than mercury) were speciated into coarse and fine particulate, which were needed by CMAQ. Exhibit D-3 and D-4 contain summaries of the profiles. Most were applied across an entire sector or multiple sectors (i.e., the nonroad profiles were applied to the nonroad-related sector and the stationary profile was applied to the stationary-related sectors). The "other mercury" profiles were SCC-specific. The spreadsheet file "nata_metals_split.xlxs" in the SupplementalData Folder contains information on how the profiles for mercury were assigned to SCCs "other mercury" categories listed in Exhibit D-4.

Also included in this appendix (not related to the source groups) are mercury speciation factors used for specific units at electricity generating units (EGUs) and for non-EGU categories (based on SCC). The same data were used as was used for the "2005 Platform –CAP-BAFM 2005-Based Platform, Version 4.1 (use for Mercury)" documented here. For EGUs, which are in the ptegu CMAQ NATA platform modeling sector, we used unit-specific speciation factors based on those developed for the Clean Air Mercury Rule (CAMR) development and documented for the Utility MACT here. These are provided in the file spreadsheet file "nata_metals_split.xlsx" in the SupplementalData folder, and the methodology used relied on matching 2011 NEI units to identifiers used in CAMR in order to use the same speciation data. New units were mapped to mercury speciation bins based on configuration. The methodology and supporting files are at "Hgunit-specific-speciation bins2011nei.zip" in the SupplementalData folder.

Exhibit D-1. Relationship of NEI Sectors with HEM Run Groups and Source Groups

NEI Data Category	EIS Sector	HEM Run Group	HEM Group Abbrev.	Source Group	Source Group Abbrev.
Point	[Multiple]	Point, excluding airports	Point	Point stationary (PT)	Point stationary (PT)
	Mobile-aircraft	Airports	Airports	NR-Railyard (PT)	NR-Railyard (PT)
	Mobile-locomotives	Point, excluding airports	Point	NR-Airport (PT)	NR-Airport (PT)
Nonpoint	Fuel Comb - Industrial Boilers, ICEs - Coal	Nonpoint 10m ReleaseHeight	NP10m	Industrial Commercial Institutional Fuel Combustion	NP-ICI_fuel_comb
	Fuel Comb - Industrial Boilers, ICEs - Oil				
	Fuel Comb - Industrial Boilers, ICEs - Natural Gas				
	Fuel Comb - Industrial Boilers, ICEs - Other				
	Fuel Comb - Industrial Boilers, ICEs - Biomass				
	Fuel Comb - Comm/Institutional - Coal				
	Fuel Comb - Comm/Institutional - Oil				
	Fuel Comb - Comm/Institutional - Natural Gas				
	Fuel Comb - Comm/Institutional - Other				
	Fuel Comb - Comm/Institutional - Biomass				
	Fuel Comb - Residential - Other	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Residential Fuel Combustion - Except Wood	NP-Non-RWC_ResFuelComb
	Fuel Comb - Residential - Oil				
	Fuel Comb - Residential - Natural Gas				
	Fuel Comb - Residential - Wood	Nonpoint Residential Wood Combustion	RWC	Nonpoint Residential Wood Combustion	RWC
Onroad	Gas Stations	Onroad Lightduty (includes refueling)	OR_LD	Onroad Refueling	OR-Refueling
	Mobile - On-Road non-Diesel Light Duty Vehicles			Onroad Lightduty Gas	OR-LightDuty_Gas
	Mobile - On-Road non-Diesel Heavy Duty Vehicles	Onroad Heavyduty	OR_HD	Onroad Heavyduty Gas	OR-HeavyDuty_Gas
	Mobile - On-Road Diesel Light Duty Vehicles	Onroad Lightduty (includes refueling)	OR_LD	Onroad Lightduty Diesel	OR-LightDuty_Diesel
	Mobile - On-Road Diesel Heavy Duty Vehicles	Onroad Heavyduty	OR_HD	Onroad Heavyduty Diesel	OR-HeavyDuty_Diesel

NEI Data Category	EIS Sector	HEM Run Group	HEM Group Abbrev.	Source Group	Source Group Abbrev.
Nonroad * (*excludes airports, CMV and locomotives)	Mobile - Non-Road Equipment - Gasoline	Nonroad * (*excludes airports, CMV and locomotives)	nonroad	Nonroad Gas Other	NR-Gas/Other
		Nonroad		Nonroad Construction	NR-Construction
	Mobile - Non-Road Equipment - Other	Nonroad		Nonroad Gas Other	NR-Gas/Other
	Mobile - Non-Road Equipment - Diesel	Nonroad		Nonroad Diesel Other	NR-Diesel/Other
				Nonroad Construction	NR-Construction
Nonpoint	Mobile - Commercial Marine Vessels	Commercial Marine Vessels (CMV)	CMV_P	Nonroad CMV Ports (shapes)	NR-CMV_ports
			CMV_UW	Nonroad CMV Underway (shapes)	NR-CMV_underway
Nonroad	Mobile - Non-Road Equipment - Gasoline	Nonroad	nonroad	Nonroad Pleasurecraft	NR-Pleasurecraft
	Mobile - Non-Road Equipment - Diesel				
Nonpoint	Mobile - Locomotives	Nonpoint Low ReleaseHeight	NPlow	Nonroad Locomotives	NR-Locomotives
	Industrial Processes - Chemical Manuf	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Chemical Manufacturing	NP-Chemical_Mfg
	Industrial Processes - NEC			Nonpoint Industrial Not Elsewhere Classified	NP-Industrial_NEC
	Commercial Cooking	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Commercial Cooking	NP-Comm_cooking
	Industrial Processes - Non-ferrous Metals	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Nonferrous Metals	NP-Nonferrous_metals
	Industrial Processes - Petroleum Refineries			Nonpoint Refineries	NP-Refineries
	Industrial Processes - Oil & Gas Production			Nonpoint Oil and Gas	NP-Oil/Gas
	Industrial Processes - Mining			Nonpoint Mining	NP-Mining
	Solvent - Non-Industrial Surface Coating	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Nonindustrial Surface Coating	NP-Non-ind_sfc_coating
	Solvent - Industrial Surface Coating & Solvent Use	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Surface Coating and Industrial Solvent	NP-SfcCoating_IndSolvent
	Solvent - Degreasing	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Degreasing	NP-Solvent_degreasing
	Solvent - Dry Cleaning				NP-Dry_cleaning
	Solvent - Graphic Arts	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Surface Coating and Industrial Solvent	NP-SfcCoating_IndSolvent
	Solvent - Consumer & Commercial Solvent Use	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Consumer Commercial Solvent	NP-Consumer_comm_solvent
	Industrial Processes - Storage and Transfer	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Storage and Transfer	NP-Storage_Transfer

NEI Data Category	EIS Sector	HEM Run Group	HEM Group Abbrev.	Source Group	Source Group Abbrev.
	Miscellaneous Non-Industrial NEC	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Miscellaneous Nonindustrial	NP-Misc_non-ind
	Bulk Gasoline Terminals	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Bulk Gasoline Terminals	NP-Bulk_gas_term
	Gas Stations	Nonpoint Low ReleaseHeight	NPlow	Nonpoint GasStations - Stage I	NP-Gas_stations
	Gas Stations				NP-Gas_stations
	Waste Disposal	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Wastedisposal Other	NP-WasteDisposal_Other
	Waste Disposal			Nonpoint Landfills	NP-Landfills
	ag burning			agburning-modeled only in CMAQ	agburning-modeled only in CMAQ
	Biogenics - Vegetation and Soil			Biogenics-modeled only in CMAQ	Biogenics-modeled only in CMAQ

Exhibit D-2 Source Group and HEM Run Group Cross-reference to Inventory SCCs

NEI Data Category	HEM Ru Group	HEM Run Group Abbrev.	Source Group	Source Group Abbrev.	SCCs
Nonpoint	Nonpoint 10m ReleaseHeight	NP10m	Industrial Commercial Institutional Fuel Combustion	NP-ICI_fuel_comb	2102001000, 2102002000, 2102004000, 2102004001, 2102004002, 2102005000, 2102006000, 2102006002, 2102007000, 2102008000, 2102011000, 2102012000, 2103001000, 2103002000, 2103004000, 2103004001, 2103004002, 2103005000, 2103006000, 2103007000, 2103008000, 2103011000, 2801520000, 2801520004, 2103010000
	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Residential Fuel Combustion - Except Wood	NP-Non-RWC_ResFuelComb	2104001000, 2104002000, 2104004000, 2104005000, 2104006000, 2104006010, 2104007000, 2104011000,
	Nonpoint Residential Wood Combustion	RWC	Nonpoint Residential Wood Combustion	RWC	2104008100, 2104008210, 2104008220, 2104008230, 2104008310, 2104008320, 2104008330, 2104008400, 2104008420, 2104008510, 2104008610, 2104008700, 2104009000,
Onroad	Onroad Lightduty (includes refueling)	OR_LD	Onroad Refueling	OR-Refueling	2201110162, 2201110262, 2201110362, 2201110462, 2201110562, 2201210162, 2201210262, 2201210362, 2201210462, 2201210562, 2201310162, 2201310262, 2201310362, 2201310462, 2201310562, 2201320162, 2201320262, 2201320362, 2201320462, 2201320562, 2201420162, 2201420262, 2201420362, 2201420462, 2201420562, 2201430162, 2201430262, 2201430362, 2201430462, 2201430562, 2201510162, 2201510262, 2201510362, 2201510462, 2201510562, 2201520162, 2201520262, 2201520362, 2201520462, 2201520562, 2201530162, 2201530262, 2201530362, 2201530462, 2201530562', 2201540162, 2201540262, 2201540362, 2201540462, 2201540562, 2201610162, 2201610262, 2201610362, 2201610462, 2201610562, 2202210162, 2202210262, 2202210362, 2202210462, 2202210562, 2202310162, 2202310262, 2202310362, 2202310462, 2202310562, 2202320162, 2202320262, 2202320362, 2202320462, 2202320562, 2202410162, 2202410262, 2202410362, 2202410462, 2202410562, 2202420162, 2202420262, 2202420362, 2202420462, 2202420562, 2202430162, 2202430262, 2202430362, 2202430462, 2202430562, 2202510162, 2202510262, 2202510362, 2202510462, 2202510562, 2202520162, 2202520262, 202520362, 2202520462, 2202520562, 2202530162, 2202530262, 2202530362', 2202530462, 2202530562, 2202540162, 2202540262, 2202540362, 2202540462, 2202540562, 2202610162, 2202610262, 2202610362, 2202610462, 2202610562, 2202620162, 2202620262, 2202620362, 2202620462, 2202620562, 2205210162, 2205210262, 2205210362, 2205210462, 2205210562, 2205310162, 2205310262, 2205310362, 2205310462, 2205310562, 2205320162, 2205320262, 2205320362, 2205320462, 2205320562
	Onroad Lightduty (includes refueling)	OR_LD	Onroad Lightduty Gas	OR-LightDuty_Gas	2201110181, 2201110281, 2201110381, 2201110481, 2201110581, 2201210181, 2201210281, 2201210381, 2201210481, 2201210581, 2201310181, 2201310281, 2201310381, 2201310481, 2201310581, 2201320181, 2201320281, 2201320381, 2201320481, 2201320581, 2205210181, 2205210281, 2205210381, 2205210481, 2205210581, 2205310181, 2205310281, 2205310381, 2205310481, 2205310581, 2205320181, 2205320281, 2205320381, 2205320481, 2205320581
	Onroad Heavyduty	OR_HD	Onroad Heavyduty Gas	OR-HeavyDuty_Gas	2201420181, 2201420281, 2201420381, 2201420481, 2201420581, 2201430181, 2201430281, 2201430381, 2201430481, 2201430581, 2201510181, 2201510281, 2201510381, 2201510481, 2201510581, 2201520181, 2201520281, 2201520381, 2201520481, 2201520581, 2201530181, 2201530281, 2201530381, 2201530481, 2201530581, 2201540181, 2201540281, 2201540381, 2201540481, 2201540581, 2201610181, 2201610281, 2201610381, 2201610481, 2201610581, 2203420181, 2203420281, 2203420381, 2203420481, 2203420581
	Onroad Lightduty (includes refueling)	OR_LD	Onroad Lightduty Diesel	OR-LightDuty_Diesel	2202210181, 2202210281, 2202210381, 2202210481, 2202210581, 2202310181, 2202310281, 2202310381, 2202310481, 2202310581, 2202320181, 2202320281, 2202320381, 2202320481, 2202320581

NEI Data Category	HEM Ru Group	HEM Run Group Abbrev.	Source Group	Source Group Abbrev.	SCCs
	Onroad Heavyduty	OR_HD	Onroad Heavyduty Diesel	OR-HeavyDuty_Diesel	2202410181, 2202410281, 2202410381, 2202410481, 2202410581, 2202420181, 2202420281, 2202420381, 2202420481, 2202420581, 2202430181, 2202430281, 2202430381, 2202430481, 2202430581, 2202510181, , 2202510281, 2202510381, 2202510481, 2202510581, 2202520181, 2202520281, 2202520381, 2202520481, 2202520581, 2202530181, 2202530281, 2202530381, 2202530481, 2202530581, 2202540181, 2202540281, 2202540381, 2202540481, 2202540581, 2202610181, 2202610281, 2202610381, 2202610481, 2202610581, 2202620153, 2202620181, 2202620191, 2202620281, 2202620381', 2202620481, 2202620581
Nonroad	Nonroad (excludes airports, CMV and locomotives)	nonroad	Nonroad Gas Other	NR-Gas/Other	2260001010, 2260001020, 2260001030, 2260001060, 22600030302, 260003040, 2260004000, 2260004015, 2260004016, 2260004020, 2260004021, 2260004025, 2260004026, 2260004030, 2260004031, 2260004035, , 2260004036, 2260004071, 2260005035, 2260006005, 2260006010, 2260006015, 2260006035, 2260007005, 2265001010, 2265001030, 2265001050, 2265001060, 2265003010, 2265003020, 2265003030, 2265003040, 2265003050, 2265003060, 2265003070, 2265004010, 2265004011, 2265004015, 2265004016, 2265004025, 2265004026, 2265004030, 2265004031, 2265004035, 2265004036', 2265004040, 2265004041, 2265004046, 2265004050, 2265004051, 2265004055, 2265004056, 2265004066, 2265004071, 2265004075, 2265004076, 2265005010, 2265005015, 2265005020, 2265005025, 2265005030, 2265005035, 2265005040, 2265005045, 2265005055, 2265005060, 2265006005, 2265006010, 2265006015, 2265006025, 2265006030, 2265006035, 2265007010, 2265007015, 2265010010, 2267001060, 2267002003, 2267002015, 2267002021, 2267002024, 2267002030, 2267002033, 2267002039, 2267002045, 2267002054, 2267002057, 2267002060, 2267002066, 2267002072, 2267002081, 2267003010, 2267003020, 2267003030, 2267003040, 2267003050, 2267003070, 2267004066, 2267005055', 2267005060, 2267006005, 2267006010, 2267006015, 2267006025, 2267006030, 2267006035, 2268002081, 2268003020, 2268003030, 2268003040, 2268003060, 2268003070, 2268005055, 2268005060, 2268006005, 2268006010, 2268006015, 2268006020, 2268006035, 2268010010, 2285004015, 2285006015
	Nonroad (excludes airports, CMV and locomotives)	nonroad	Nonroad Construction	NR-Construction	2260002000, 2260002006, 2260002009, 2260002021, 2260002027, 2260002039, 2260002054, 2265002003, 2265002006, 2265002009, 2265002015, 2265002021, 2265002024, 2265002027, 2265002030, 2265002033, 2265002039, 2265002042, 2265002045, 2265002054, 2265002057, 2265002060, 2265002066, 2265002072, 2265002078, 2265002081, 2270002003, 2270002006, 2270002009, 2270002015, 2270002018, 2270002021, 2270002024, 2270002027, 2270002030, 2270002033, 2270002036, 2270002039, 2270002042, 2270002045, 2270002048, 2270002051, 2270002054, 2270002057, 2270002060', 2270002066, 2270002069, 2270002072, 2270002075, 2270002078, 2270002081, 2270009010
	Nonroad (excludes airports, CMV and locomotives)	nonroad	Nonroad Diesel Other	NR-Diesel/Other	2270001060, 2270003010, 2270003020, 2270003030, 2270003040, 2270003050, 2270003060, 2270003070, 2270004031, 2270004036, 2270004046, 2270004056, 2270004066, 2270004071, 2270004076, 2270005010, , 2270005015, 2270005020, 2270005025, 2270005030, 2270005035, 2270005040, 2270005045, 2270005055, 2270005060, 2270006005, 2270006010, 2270006015, 2270006020, 2270006025, 2270006030, 2270006035, 2270007010, 2270007015, 2270010010, 2285002015
Nonpoint	Commercial Marine Vessels (CMV)	CMV_P	Nonroad CMV Ports (shapes)	NP-CMV_ports	2280002100, 2280003100
		CMV_UW	Nonroad CMV Underway (shapes)	NP-CMV_underway	2280002200, 2280003200
Nonroad	Nonroad (excludes airports, CMV and locomotives)	nonroad	Nonroad Pleasurecraft	NR-Pleasurecraft	2282005010, 2282005015, 2282010005, 2282020005, 2282020010

NEI Data Category	HEM Ru Group	HEM Run Group Abbrev.	Source Group	Source Group Abbrev.	SCCs
Nonpoint	Nonpoint Low ReleaseHeight	NPlow	Nonroad Locomotives	NP-Locomotives	2285002006, 2285002007, 2285002008, 2285002009, 2285002010
	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Chemical Manufacturing	NP-Chemical_Mfg	2301000000, 2301020000,
			Nonpoint Industrial Not Elsewhere Classified	NP-Industrial_NEC	2302000000, 2302080000, 2305000000, 2305070000, 2305080000, 2307000000, 2308000000, 2309000000, 2309100010, 2309100030, 2309100050, 2312000000, 2399000000
	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Commercial Cooking	NP-Comm_cooking	2302002000, 2302002100, 2302002200, 2302003000, 2302003100, 2302003200
	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Nonferrous Metals	NP-Nonferrous_metals	2304000000, 2303000000
			Nonpoint Refineries	NP-Refineries	2306000000, 2306010000
			Nonpoint Oil and Gas	NP-Oil/Gas	2310000000, 2310000220, 2310000230, 2310000330, 2310000550, 2310000660, 2310010000, 2310010100, 2310010200, 2310010300, 2310011000, 2310011201, 2310011501, 2310011502, 2310011503, 2310011505, , 2310020000, 2310020600, 2310021010, 2310021030, 2310021100, 2310021202, 2310021209, 2310021251, 2310021300, 2310021302, 2310021309, 2310021310, 2310021351, 2310021400, 2310021500, 2310021501, 2310021502, 2310021503, 2310021505, 2310021506, 2310021509, 2310021603, 2310021700, 2310030000, 2310111100, 2310111401, 2310111700, 2310121100, 2310121401', 2310121700, 2310002401, 2310002411, 2310002421, 2310010700, 2310010800, 2310011020, 2310011100, 2310011450, 2310011500, 2310011504, 2310011506, 2310011600, 2310012020, 2310012511, 2310012512, 2310012515, 2310012516, 2310012521, 2310012522, 2310012526, 2310020700, 2310020800, , 2310021101, 2310021102, 2310021103, 2310021201, 2310021203, 2310021301, 2310021303, 2310021504, 2310021600, 2310021601, 2310021602, 2310021604, 2310021605, 2310022010, 2310022090, 2310022105, 2310022420, 2310022501, 2310022502, 2310022505, 2310022506, 2310023010, 2310023030, 2310023100, 2310023102, 2310023202, 2310023251, 2310023300, 2310023302, 2310023310', 2310023351, 2310023400, 2310023509, 2310023511, 2310023512, 2310023513, 2310023515, 2310023516, 2310023600, 2310023601, 2310023602, 2310023603, 2310023606, 2310030210, 2310030300, 2310030401, 2310112401, 2310122100
			Nonpoint Mining	NP-Mining	2325030000, 2325060000
	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Nonindustrial Surface Coating	NP-Non-ind_sfc_coating	2401001000, 2401002000, 2401003000
	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Surface Coating and Industrial Solvent	NP-SfcCoating_IndSolvent	2401005000, 2401005500, 2401005600, 2401005700, 2401005800, 2401008000, 2401010000, 2401015000, 2401020000, 2401025000, 2401030000, 2401035000, 2401040000, 2401045000, 2401050000, 2401055000, , 2401060000, 2401065000, 2401070000, 2401075000, 2401080000, 2401085000, 2401090000, 2401100000, 2401200000, 2425000000, 2425010000, 2425020000, 2425030000, 2425040000, 2440000000, 2440020000, 2402000000
	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Degreasing	NP-Solvent_degreasing	2415000000, 2415005000, 2415010000, 2415020000, 2415025000, 2415030000, 2415035000, 2415040000, 2415045000, 2415050000, 2415055000, 2415060000, 2415065000, 2415100000, 2415130000, 2415230000, , 2415245000, 2415300000, 2415345000, 2415360000

NEI Data Category	HEM Ru Group	HEM Run Group Abbrev.	Source Group	Source Group Abbrev.	SCCs
			Nonpoint Drycleaning	NP-Dry_cleaning	2420000000, 2420000055, 2420010000, 2420010055, 2420010370, 2420020000
			Nonpoint Consumer Commercial Solvent	NP-Consumer_comm_solvent	2460000000, 2460100000, 2460110000, 2460120000, 2460130000, 2460150000, 2460160000, 2460170000, 2460180000, 2460190000, 2460200000, 2460210000, 2460220000, 2460230000, 2460250000, 2460270000, , 2460290000, 2460400000, 2460410000, 2460420000, 2460500000, 2460510000, 2460520000, 2460600000, 2460610000, 2460800000, 2460810000, 2460820000, 2460900000, 2461021000, 2461022000, 2461023000, 2461800000, 2461850000, 2461850001, 2461850004, 2461850005, 2461850006, 2461850009, 2461850051, 2461850052, 2461850053, 2461850054, 2461850055, 2461850056', 2461850099, 2465000000, 2465100000, 2465200000, 2465400000, 2465800000
	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Storage and Transfer	NP-Storage_Transfer	2501000150, 2501995120, 2501995150, 2505000120, 2505010000, 2505020000, 2505020030, 2505020060, 2505020090, 2505020093, 2505020120, 2505020121, 2505020150, 2505020180, 2505030120, 2505030150, , 2505040120, 2510010000, 2520010000
	Nonpoint Low ReleaseHeight	NPlow	Nonpoint Miscellaneous Nonindustrial	NP-Misc_non-ind	2501011011, 2501011012, 2501011013, 2501011014, 2501011015, 2501012011, 2501012012, 2501012013, 2501012014, 2501012015, 2810025000, 2810030000, 2810035000, 2810050000, 2810060100, 2810060200, , 2840000000, 2840010000, 2850000000, 2850000010, 2850001000, 2851001000, 2861000000, 2861000010, 2862000000
	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Bulk Gasoline Terminals	NP-Bulk_gas_term	2501050120, 2501055120
	Nonpoint Low ReleaseHeight	NPlow	Nonpoint GasStations - Stage I	NP-Gas_stations	2501060050, 2501060051, 2501060052, 2501060053, 2501060100, 2501060101, 2501060102, 2501060103, 2501060200, 2501060201, 2501070100, 2501080050, 2501080100, 2501080201
	Nonpoint 10m ReleaseHeight	NP10m	Nonpoint Wastedisposal Other	NP-WasteDisposal_Other	2601000000, 2610000100, 2610000300, 2610000400, 2610000500, 2610030000, 2630000000, 2630020000, 2630020020, 2635000000, 2640000000, 2660000000, 2680001000, 2650000000, 2650000002
			Nonpoint Landfills	NP-Landfills	2620000000, 2620030000, 2620030001
			agburning-modeled only in CMAQ	agburning-modeled only in CMAQ	2801500000, 2801500100, 2801500141, 2801500150, 2801500170, 2801500181, 2801500220, 2801500250, 2801500261, 2801500262, 2801500300, 2801500320, 2801500330, 2801500350, 2801500390, 2801500410, , 2801500420, 2801500430, 2801500440, 2801500450, 2801500500, 2801500600
			Biogenics-modeled only in CMAQ	Biogenics-modeled only in CMAQ	2701200000

Exhibit D-3. Speciation of Xylenes and Metals

Xylenes					
Source Type	**Profile**	**Pollutant**	**MXYL**	**OXYL**	**PXYL**
Other	0000	XYLS	0.52	0.16	0.16
Nonroad	0001	XYLS	0.68	0.32	
Onroad	0002	XYLS	0.74	0.26	

Metals				
Source Type	**Profile**	**Metal**	**Fine**	**Coarse**
Nonroad	NOARS	ARSENIC	0.83	0.17
	NONBE	BERYLLIUM	0.39	0.61
	NONCD	CADMIUM	0.38	0.62
	NONMN	MANGANESE	0.67	0.33
	NONNI	NICKEL	0.49	0.51
	NONPB	LEAD	0.88	0.12
Stationary	STANI	NICKEL	0.59	0.41
	STACD	CADMIUM	0.76	0.24
	STAMN	MANGANESE	0.67	0.33
	STAPB	LEAD	0.74	0.26
	STABE	BERYLLIUM	0.68	0.32
	CRSTA	CHROMHEX	0.71	0.29
	STARS	ARSENIC	0.59	0.41
Onroad		CHROMHEX	0.86	0.14
		MANGANESE	0.64	0.36
		NICKEL	0.83	0.17
		ARSENIC	0.95	0.05

Exhibit D-4. Speciation of Mercury (Other than EGUs)

Mobile Mercury					
Mobile	**Profile**	**Pollutant**	**HGNRVA**	**HGIIGAS**	**PHGI**
Mobile Diesel	HGMD	EXH__HGSUM	0.56	0.29	0.15
Mobile Gasoline	HGMG	EXH__HGSUM	0.91	0.086	0.004

Other Mercury				
Profile Code	**Description**	**Elemental**	**Divalent Gas**	**Particulate**
HBCMB	combustion	0.5	0.3	0.2
HGCEM	cement	0.75	0.13	0.12
HGCHL	chloralkali processes	0.95	0.05	0
HGGLD	Gold mining	1	0	0
HGINC	Incineration	0.22	0.58	0.2
HGMD	Mobile diesel	0.56	0.29	0.15
HGMG	Mobile gas	0.91	0.086	0.004
HGIND	Other Industrial	0.8	0.1	0.1

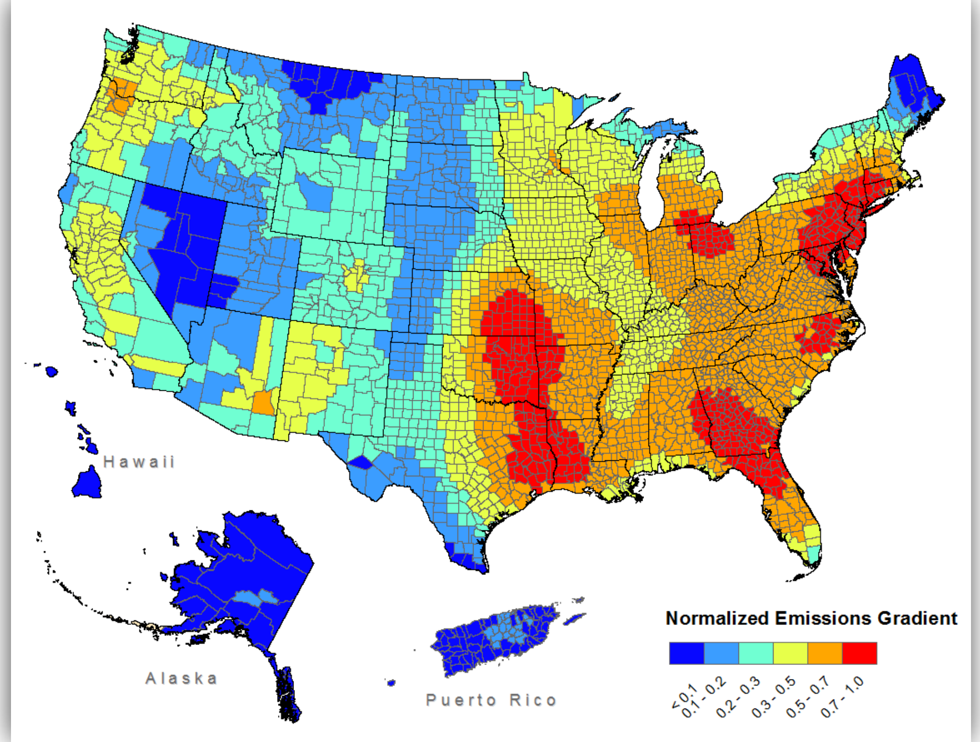
Note: EGUs use unit-specific profiles provided in the SupplementalData Folder spreadsheet file “nata_metals_split.xlsx”.

Appendix E

Estimation of Background Concentrations for the 2011 NATA

The report in this appendix contains the methods we used to estimate background concentrations for NATA.

Methods for Estimating Background Concentrations for the National Air Toxics Assessment (NATA) 2011

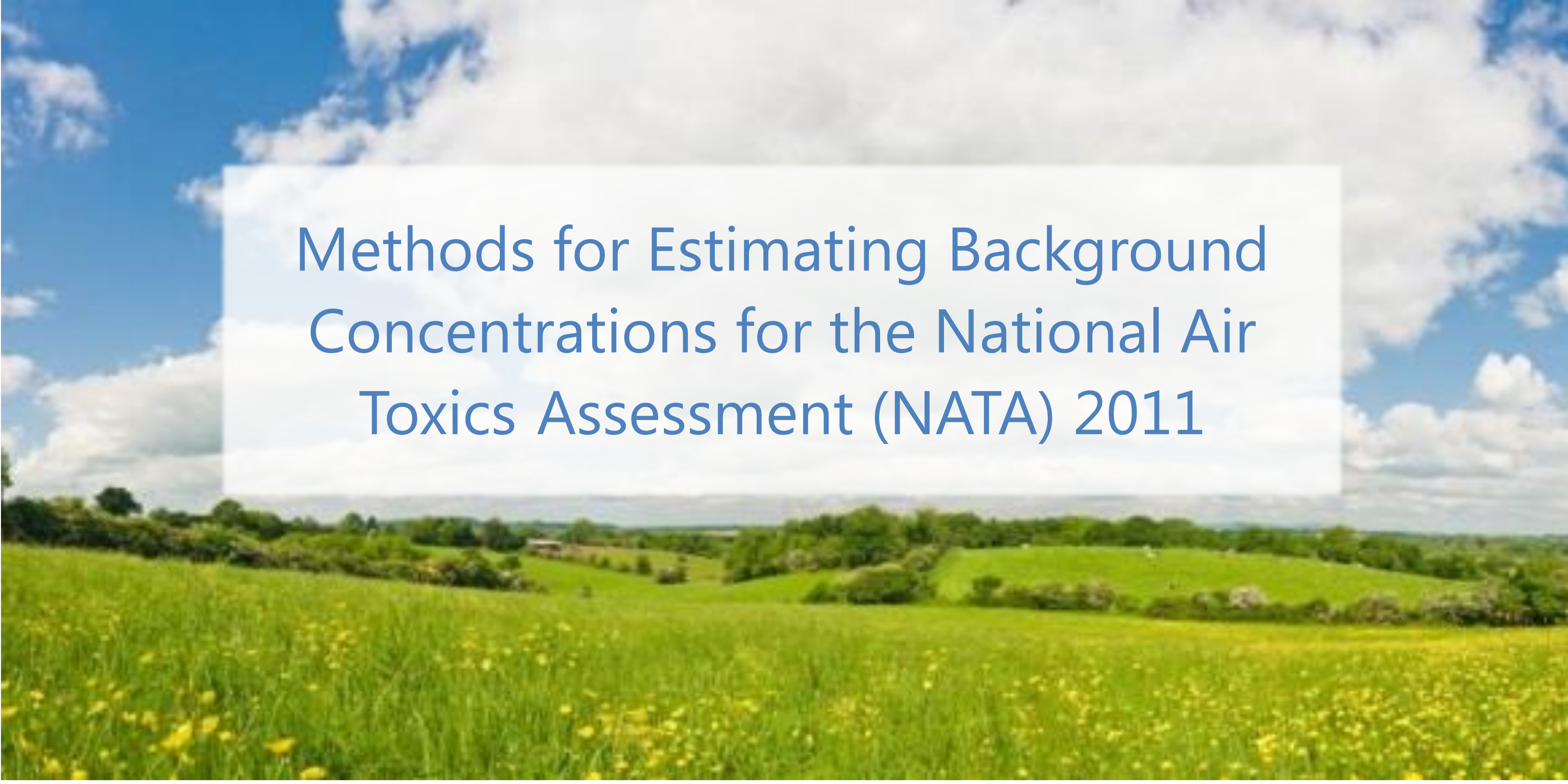


Final Report Prepared for

U.S. Environmental Protection Agency
Research Triangle Park, NC

August 2015

STi
Sonoma Technology, Inc.



Methods for Estimating Background Concentrations for the National Air Toxics Assessment (NATA) 2011

Prepared by

Michael McCarthy, PhD
Theresa O’Brien
Yuan Du
Ashley Russell, PhD

Sonoma Technology, Inc.
1455 N. McDowell Blvd., Suite D
Petaluma, CA 94954-6503
Ph 707.665.9900 | F 707.665.9800
sonomatech.com

Prepared for

Dennis Doll
Madeleine Strum
Ted Palma

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
109 T.W. Alexander Dr.
Research Triangle Park, NC 27709
919.541.5693

Final Report

STI-915110-6315

August 13, 2015

Cover graphic illustrates the normalized emissions gradient for benzene. See Section 4.2 for details.

Contents

Figures

Tables

1. Introduction

1.1 Overview

This document describes the methods used to estimate "background" concentrations for the U.S. Environmental Protection Agency's (EPA) 2011 National Air Toxics Assessment (NATA). Background concentrations were derived for the previous NATA modeling estimates in 2005, 2002, 1999, and 1996. In most previous iterations of NATA, "background" concentrations were defined as that portion of concentrations reflecting contributions transported from farther than 50 km, emissions originating from outside the United States, unidentified emissions sources within a 50-km buffer, and natural emissions sources. These contributions can be significant for some air toxics. In the 2005 NATA, the chemical transport model CMAQ (Community Multi-scale Air Quality model) was used for a subset of pollutants; the CMAQ model eliminates the need to account for emissions originating from beyond 50 km and outside the model domain when examining background concentrations.

For the 2011 NATA, the background concentration estimates will be applied pollutant by pollutant, and only in areas outside the contiguous states where the CMAQ model domain does not extend. (The final NATA 2011 documentation will describe exactly which background concentrations were applied for each pollutant in each area of the NATA domain.) In other words, many of the background estimates described in this report will be overridden in NATA 2011 by the CMAQ predictions. Thus, we caution that results shown here should not be considered representative of the background concentrations applied in NATA 2011.

1.2 Background Definitions

For the 2011 NATA, we investigated two types of "background" concentration estimates:

- **Remote Northern Hemisphere (NH) background.** The annual mean concentration at remote receptor areas not impacted by local-scale (50 km) or regional-scale emissions of the pollutant. Examples of remote locations include the summit of Mauna Loa, the Aleutian Islands in Alaska, portions of the Pacific Coast between Oregon and California, and some parts of the Mountain West. These estimates represent the lowest concentrations that would be observed in the United States. Generating remote background concentrations is the goal of this project.
- **Regional background.** The annual mean concentration of air in locations uninfluenced by local-scale (50 km) emissions sources of the pollutant. These background concentrations are important for densely populated sections of the country, such as the industrial Midwest, the East Coast, and portions of the Southeast. Regional background concentrations are higher than remote NH background concentrations and capture the regional transport of emissions

> over scales of hundreds of kilometers. Generating reasonable regional background concentrations was the operational goal of previous NATA iterations.

Ultimately, our investigation did not demonstrate that regional background concentration estimates were statistically reliable, and thus those concentrations were not applied in the 2011 NATA. See Section 4.2 for additional details on the regional background issues.

1.3 NATA Background Approaches

In previous NATA iterations, different approaches have been applied to develop background concentrations. In the 1996 effort, background concentrations were gathered in a literature search performed as part of the Cumulative Exposure Project (CEP). The CEP literature review was originally performed to acquire background concentrations for 1990 (Rosenbaum et al., 1999; Woodruff et al., 1998). The result of the literature search was a single remote background value representing 12 air toxics.

For the 1999 NATA, two approaches were used to estimate background concentrations (Bortnick et al., 2003). The primary approach used measurements from ambient monitors to estimate background concentrations. Estimates from individual locations were extrapolated to counties without measurements and were based on a population regression. When ambient measurements were not available from the ambient monitoring network, background concentrations from the CEP were used. **Table 1** lists the pollutants for which background concentrations were estimated for NATA 1996 and 1999.

For the 2002 and 2005 NATAs, background concentrations were developed using a three-pronged approach that selected background concentrations depending on the atmospheric residence times of the pollutants, an ambient network method, and an emissions-based method. In this approach, background concentrations were assigned for four chemicals with known, globally averaged concentrations (i.e., long-residence times). Background concentrations were determined for another 13 pollutants by a method that utilized routine ambient measurements. Finally, background concentrations for another 15 pollutants were determined using a novel, emissions-based method. Each of these methods is described in McCarthy et al. (2008). **Table 2** lists the pollutants and approach used for the NATA 2002 and 2005 results.

For the 2011 NATA, Sonoma Technology, Inc. (STI) developed spatially uniform estimates of remote Northern Hemisphere background concentrations for each of the air toxics of interest, following previous approaches with some modifications. In addition, STI investigated emissions-based and ambient-based methods for estimating spatially varying background concentrations. However, because of a lack of agreement and predictive capacity between these two methods, they were not applied in the final analysis.

Section 2 of this report describes the technical approach used, Section 3 presents results, and Section 4 discusses their implications for contributing to modeled cancer risk and hazard. Section 5 briefly summarizes the results, and Section 6 lists references cited in the report.

Table 1. Pollutants for which background concentrations were estimated in the NATA 1996 and 1999 modeling efforts. This table is adapted from the NATA 1999 website.[1]

NATA 1996	NATA 1999	
Benzene	Benzene	1,3-Butadiene
Carbon tetrachloride	Carbon tetrachloride	Bis(2-ethylhexyl)phthalate
Chloroform	Chloroform	Bromoform
Dichloromethane (methylene chloride)	Dichloromethane	Carbon disulfide
Ethylene dibromide (1,2-dibromoethane)	Ethylene dibromide	Chlordane
Ethylene dichloride (1,2-dichloroethane)	Ethylene dichloride	Hexachlorobutadiene
Formaldehyde	Formaldehyde	Hexachloroethane
Mercury	Mercury	Lindane
Polychlorinated biphenyls	Polychlorinated biphenyls	Methyl bromide (bromomethane)
Tetrachloroethene (perchloroethylene, tetrachloroethylene)	Tetrachloroethene	Methyl chloride (chloromethane)
Trichloroethene (trichloroethylene)	Trichloroethene	Methyl chloroform (1,1,1-trichloroethane)
Hexachlorobenzene	Acetaldehyde	Phosgene
	1,1,2,2-Tetrachloroethane	Vinyl chloride
	1,2-Dichloropropane (propylene dichloride)	Xylenes

---

[1] http://www.epa.gov/ttn/atw/nata1999/99pdfs/backgroundtable.pdf.

Table 2. Hazardous air pollutants (HAPs) and methods applied to determine background concentrations for NATA 2002 and 2005. Formaldehyde and acetaldehyde concentrations (in italics) were determined using the Community Multiscale Air Quality (CMAQ) model for the NATA 2005 exercise.

Globally-Averaged Concentrations (Long Residence Time)	Ambient-Based Method	Emissions-Based Method
Carbon tetrachloride	1,3-Butadiene	Hydrazine
Chloromethane (methyl chloride)	1,4-Dichlorobenzene (p-dichlorobenzene)	Chromium (VI)
Bromomethane (methyl bromide)	*Acetaldehyde*	Ethylene dichloride (1,2-dichloroethane)
Methyl chloroform (1,1,1-trichloroethane)	Arsenic	Naphthalene
	Benzene	1,2-Dichloropropane (propylene dichloride)
	Chloroform	Ethylene oxide
	Chromium	Acrylonitrile
	Dichloromethane (methylene chloride)	Cadmium
	Formaldehyde	Beryllium
	Lead	Ethylene dibromide (1,2-dibromomethane)
	Manganese	Benzidine
	Nickel	Quinoline
	Tetrachloroethene (perchloroethylene, tetrachloroethylene)	Bis(2-ethylhexyl)phthalate
	Toluene	1,2-Dibromo-3-chloropropane
		Trichloroethene (trichloroethylene)
		Vinyl chloride
		1,1,2,2-Tetrachloroethane

2. Technical Approach

The project's initial goal was to develop spatially varying background concentrations appropriate to individual county-level estimates for the air toxics of interest. In the initial approach, the remote-concentration methodology was only a step to provide a lowest-level concentration estimate for air toxics. The lowest-level concentration would bind the lowest possible county concentrations for remote areas of the United States uninfluenced by regional emission sources. Regional influences would be estimated by examining emissions gradients and ambient concentrations measured at sites across the country. Counties without measurements would be assigned background concentration estimates according to a regression of ambient data against predicted emissions gradients.

As described in Section 4.2, using the combined emissions-based and ambient-based approach to predict spatial variability in background concentrations showed no statistically significant relationship between the predicted ambient and emissions-gradient background concentrations for any of the primary pollutants of interest. Therefore, for the 2011 NATA background concentration estimates, we have chosen to use only remote concentration estimates (RCEs).

2.1 Methods for Remote Concentration Estimates

Multiple methods were used to develop estimates of remote Northern Hemisphere background air toxics concentrations for NATA 2011. These methods include using concentrations reported in the scientific literature, averages of measurements made at remote monitoring sites, estimates based on national monitoring network lower-level concentrations, and estimates based on national emissions and atmospheric residence times. The generalized hierarchy was as follows:

1. **Remote network method (Section 2.1.1).** Measurements made at networks/sites in remote Northern Hemisphere locations with citations in peer-reviewed literature. Examples include National Oceanic and Atmospheric Administration (NOAA) Global Monitoring Division (GMD) sites,[2] the Trinidad Head Advanced Global Atmospheric Gases Experiment (AGAGE) site,[3] and remote Interagency Monitoring of Protected Visual Environments (IMPROVE) sites.[4]
2. **Ambient national network method (Section 2.1.2).** Estimates of background concentrations made at routine monitoring network sites in the United States where concentrations were measurable and reliable down to the 10th percentile.
3. **Literature-based method (Section 2.1.3).** Measurements of species at remote or regional sites during a single measurement study for a specific pollutant. These measurements are different from the ongoing measurements used in Step 1 because of their “one-off” nature.

---

[2] www.esrl.noaa.gov/gmd/.

[3] http://agage.mit.edu/. Other AGAGE sites are available, but they are not upwind of the continental U.S.

[4] http://vista.cira.colostate.edu/improve/.

4. **Emissions-based method (Section 2.1.4).** Emissions-based estimates of remote background concentrations. These estimates are derived from the 2011 National Emission Inventory, atmospheric residence times, and a comparison to tetrachloroethene.

The first method relies on routine monitoring networks dedicated to measurements at remote locations that are likely to have minimal influence from local emissions. When available, these measurements were considered the best and most reliable source of remote concentrations for the year 2011. However, these networks do not target the entire list of air toxics of concern. In addition, some of the metals measured at the IMPROVE monitoring sites have method detection limits (MDLs) that are too high to accurately constrain the remote concentrations for some of their target pollutants.

The second method relies on using ambient measurements taken from 2010-2012 at routine monitoring sites in the United States. These sites are often located in urban areas and may be affected by local and regional emissions. However, the lower concentrations (i.e., 10th percentile) observed at these sites may be representative of the transported regional background concentrations. If measured reliably, these measurements were considered a viable means of estimating background concentrations.

The third method uses literature estimates from single studies that measured a pollutant of interest. This method was applied when ambient measurements were too high for estimating remote concentrations.

Finally, when none of the above three measurement-based methods produced a value, an emissions-based method was applied. Table 3 lists the pollutants for which remote concentration estimates were developed and shows the method selected to generate the remote concentration value. For some pollutants, multiple methods were applied; in such cases, the generalized hierarchy was used to select the most appropriate background concentration estimate. Only two pollutants from the 2002 and 2005 NATA background lists were excluded: quinoline and total chromium (though hexavalent chromium is included in 2011 NATA). Four pollutants (or grouped pollutants) that had been in the 1999 NATA background list were excluded: PCBs, lindane, phosgene, and hexachlorobutadiene. In contrast, 17 pollutants and groups that had never previously been estimated were added. These include acetonitrile, acrolein, hexane, 2,2,4-trimethylpentane, ethylbenzene, methyl isobutyl ketone, styrene, 1,3-dichloropropene, 1,1,2-trichloroethane, isopropylbenzene, selenium, propanal, antimony, cobalt, and vinyl acetate. As noted in the introduction, estimates developed here may not be used for some of the pollutants or in certain locations; see the final NATA documentation for a description of the relevant background values applied in the NATA 2011.

Table 3. List of pollutants and method used to generate remote concentration estimates.

Pollutant	Method[a]	Pollutant	Method[a]
1,1,2,2-Tetrachloroethane	Emissions	1,3-Butadiene	Literature
1,1,2-Trichloroethane	Emissions	Acetaldehyde	Literature
1,2-Dibromo-3-chloropropane	Emissions	Acetonitrile	Literature
1,2-Dichloropropane (propylene dichloride)	Emissions	Acrolein	Literature
1,3-Dichloropropene	Emissions	Bromoform	Literature
1,4-Dichlorobenzene (p-dichlorobenzene)	Emissions	Ethylbenzene	Literature
2,2,4-Trimethylpentane (iso-octane)	Emissions	Formaldehyde	Literature
Acrylonitrile	Emissions	Toluene	Literature
Antimony	Emissions	Xylenes	Literature
Benzidine	Emissions	Arsenic	Ambient national
Benzyl chloride	Emissions	Beryllium	Ambient national
Bis(2-ethylhexyl) phthalate	Emissions	Cadmium	Ambient national
Carbon disulfide	Emissions	Chromium VI	Ambient national
Ethylene dibromide (1,2-dibromoethane)	Emissions	Cobalt	Ambient national
Ethylene dichloride (1,2-dichloroethane)	Emissions	Selenium	Ambient national
Ethylene oxide	Emissions	Benzene	Remote network
Hydrazine	Emissions	Carbon tetrachloride	Remote network
Isopropylbenzene (cumene)	Emissions	Chloroform	Remote network
Methyl isobutyl ketone	Emissions	Chloromethane (methyl chloride)	Remote network
Naphthalene	Emissions	Dichloromethane (methylene chloride)	Remote network
n-Hexane	Emissions	Lead	Remote network

Pollutant	Method[a]	Pollutant	Method[a]
PAHs (polycyclic organic matter – POM)[b]	Emissions	Manganese	Remote network
Propanal (propionaldehyde)	Emissions	Mercury (gaseous)	Remote network
Styrene	Emissions	Methyl bromide (bromomethane)	Remote network
Vinyl acetate	Emissions	Methyl chloroform (1,1,1-trichloroethane)	Remote network
Vinyl chloride	Emissions	Nickel	Remote network
		Tetrachloroethene (perchloroethylene, tetrachloroethylene)	Remote network
		Trichloroethene (trichloroethylene)	Remote network

[a] **Remote network:** generated by monitoring networks with sites in remote areas.
Ambient national: generated by national networks in the United States.
Literature: based on a single literature study in a remote location.
Emissions: based on emissions-based method.
[b] Estimates are not used in NATA because of uncertainty.

2.1.1 Remote Network Method

STI used concentrations from networks with monitoring sites in remote locations, where little influence from regional emissions is expected, to estimate background concentrations for some of the most important (from a risk perspective) air toxics. Remote NOAA GMD sites[5] and the Trinidad Head AGAGE site[6] primarily measure gases, and remote IMPROVE sites[7] measure metals.

Annual mean concentrations for 2011 were generated at five NOAA GMD sites: Cape Kumukahi, HI (KUM); Mauna Loa, HI (MLO); Niwot Ridge, CO (NWR); Barrow, AK (BRW); and Alert, Canada (ALT). These annual mean concentrations were cosine-weighted by latitude to generate an annual mean background estimate for carbon tetrachloride, benzene, dichloromethane, tetrachloroethene, and bromomethane. Measurement methods from this network have been reported in peer-reviewed articles such as Montzka et al. (2011; 1999). Cosine-weighted latitude averaging accounts for differences in the amount of northern hemisphere air (i.e., there is more air at the equator then at the poles).

---

[5] www.esrl.noaa.gov/gmd/.
[6] http://agage.mit.edu/.
[7] http://vista.cira.colostate.edu/improve/.

Annual mean concentrations for 2011 of chloromethane, trichloroethene, and chloroform were measured at AGAGE network sites. While multiple AGAGE sites are available in remote locations such as Mace Head, Ireland, and Cape Grim, Tasmania, the only measurements representative of the eastern Pacific Ocean are made at Trinidad Head, California; data from this site were used to generate remote concentration estimates. AGAGE data can be accessed at http://cdiac.ornl.gov/ndps/alegage.html (DOI: 10.3334/CDIAC/atg.db1001), and the original reference for the network is available from Prinn et al. (2000).

Metals are measured at the IMPROVE sites across the U.S. Some of these sites are representative of clean air coming off the Pacific Ocean and were used to generate remote background concentration estimates for lead, manganese, and nickel. Annual mean concentrations from 2010 to 2012 were generated for data from the Denali, Alaska (DENA); Kalmiopsis, Oregon (KALM); Point Reyes, California (PORE); Redwoods, California (REDW); Trapper Creek, Alaska (TRCR); Tuxedni, Alaska (TUXE); and the Haleakala, Hawaii (HACR) sites. These were cosine-latitude averaged to generate mean concentrations for the toxics measurements made at those sites. The IMPROVE network data can be accessed at http://vista.cira.colostate.edu/improve/default.htm, and recent work was described by Hyslop and White (2011).

In each case, the remote concentration estimates made from these networks were corroborated using the next three methods as well. If a reliable lower estimate could be generated using the ambient or literature methods, it supplanted this primary network as the estimate of choice. For example, arsenic was estimated using the IMPROVE network data, but slightly lower concentration estimates were generated using the ambient network method described in Section 2.1.2.

2.1.2 Ambient National Network Method

Ambient air toxics data were acquired for 2010 through 2012 from the EPA's Air Toxics Monitoring Archive (AMA)[8]. Data from AMA were acquired in February 2014. Air toxics measurements are primarily collected as 24-hr duration samples. These samples are most often collected at 1-in-3-, 1-in-6-, or 1-in-12-day frequencies. Any samples collected with less than 24-hr duration (e.g., 1-hr or 3-hr samples) were aggregated into 24-hr averages if measurements were collected for at least 75% of the day in the AMA. For example, at least 18 1-hr samples were required for aggregation to a 24-hr average. This criterion ensured reasonable diurnal concentration representation. In addition, daily data were adjusted to local conditions (LC) if they were reported in units of standard temperature and pressure (STP) using local pressure and temperature conditions in the AMA.

The following steps were initially developed with the intention of providing spatially varying background concentrations. However, given the results discussed in Section 4, we converted this methodology into an independent method for assessing remote background concentrations. The chromium VI remote concentration estimate (RCE) was based on a hybrid approach in which

---

[8] EPA, 2013. Ambient Monitoring Archive for HAPs, Phase VII. Prepared by Eastern Research Group, Inc. Delivered February 28, 2013.

concentrations of Chromium $PM_{2.5}$ from the IMPROVE remote sites and the ratio of Cr VI:Total Cr was calculated at ambient network sites. This will be discussed in more detail in Section 3.1.

Remote background concentration estimates were then developed using the following ambient-based method. This method consists of seven general steps:

1. Determine annual completeness for each parameter at each site to screen out those parameters at sites with incomplete (i.e., unrepresentative) years. For NATTS program measurements, completeness was based on completeness for a given parameter; for other programs, completeness was based on the number of daily average samples employing the same method. In both cases, require 11 daily average samples per calendar quarter and 3 valid calendar quarters. Require at least one valid year from 2010-2012 for inclusion.
2. Isolate data from valid years into year-seasons based on warm (April to September) and cold seasons (October through March).
3. Calculate 10th percentile concentration for each valid year-season.
4. Is the average MDL greater than the RCE (developed from the remote, literature, or emissions methods, in preferential order)?
   a. If yes, is the 10th percentile greater than the RCE?
      - If yes, use the 10th percentile for seasonal averaging.
      - If no (which means ambient data should not be used), use RCE for seasonal averaging.

   b. If no, is the 10th percentile greater than the MDL?
      - If no (which means ambient data must be less than RCE from other methods and are at MDL or lower), use the MDL for seasonal averaging and add flag for RCE QC check.
      - If yes, use the 10th percentile for seasonal averaging.
        - If the 10th percentile is less than or equal to the RCE, add flag for RCE QC check.
        - If the 10th percentile is greater than the RCE, no flag is needed.
5. Average the warm and cold season 10th percentiles for 2010-2012 to get the overall 10th percentile estimate for each site. If there is more than one 10th percentile estimate for a site-parameter (due to collocated measurements or method changes), then
   a. Use the estimate with the lowest average MDL (i.e., most sensitive method).
   b. If the MDLs are the same, use the estimate with the most measurements.
   c. If the sample counts are the same, then average the estimates.
6. Select the lowest background estimate from each county for each parameter.
7. For those parameters measured in at least 18 counties, compare the concentration estimates from each county to the best remote concentration estimate method for the alternate methods (remote network, literature, and emissions-based method). If at least four county estimates are below the other remote concentration estimate, assess the average percent difference between the ambient county estimates and the other remote concentration

estimates. Choose the ambient-based approach if the average percent concentration difference is lower than alternate methods by more than 30%. Average the concentrations at all sites below RCE to generate an ambient based RCE. The ambient network RCE is used only when it is lower than other methods; the ambient network RCE is not used as the primary estimate when it is higher, since it is more likely to be influenced by regional emissions.

2.1.3 Literature-Based Method

A literature search for remote concentration estimates was performed for air toxics on the target list. Two citation-based methods were applied. In the first method, a primary publication (e.g., Rosenbaum et al., 1999; Woodruff et al., 1998; McCarthy et al., 2006) was investigated for follow-up citations. Any citations that cited these primary documents were then followed to examine whether other references in their bibliography provided useful, updated remote concentration estimates for any hazardous air pollutants of concern. In the second method, keyword searches were performed for each of the key target air toxics that were expected to have potential contributions to risk or hazard based on NATA 2005 risk driver classifications.

2.1.4 Emissions-Based Method

Almost half of the air toxics listed in Table 3 were not measured in remote monitoring networks, had inadequate ambient data for the ambient-based method, and were not identified in literature searches for remote background concentration estimates. Remote background estimates based on the available ambient data for these pollutants would either be represented by too few sites from which to extrapolate data or represent poor quality measurements (i.e., mostly below MDL).

The minimum ambient concentration represents a geographically remote concentration estimate. Remote concentration estimates were based on (1) the sum of 2011 v2.0 National Emissions Inventory (NEI) emissions[9], (2) atmospheric residence times gleaned from the literature, and (3) comparison to a long-lived pollutant with measured remote concentrations and substantial U.S. emissions. In this and previous studies, that pollutant is tetrachloroethene (C_2Cl_4). **Equation 1** shows the relationship used to derive these remote estimates:

$$[C_i] = \frac{E_i * t_i * [C_{tetrachloroethene}]}{E_{tetrachloroethene} * t_{tetrachloroethene_i}} \qquad (1)$$

where *[C]* is the remote concentration, *E* is the 2011 NEI value in tons per year, *t* is the residence time in years, and *i* is the pollutant of interest. Most residence time estimates were developed for NATA 2005 and were directly used in this work.

---

[9] NEI v2.0 was a pre-released version from September 22, 2014 for all emissions categories other than onroad mobile. Onroad mobile was not available; v1.0 onroad mobile emissions were used in its place.

3. Results

Sections 3.1–3.4 present background concentrations for all the pollutants listed in Table 3.

3.1 Remote Network Estimates

Estimates of Northern Hemisphere concentrations from the remote network are presented in **Table 4**. For the IMPROVE sites, annual mean concentrations were generated for each parameter for 2010-2012. Each site was then averaged across all three years. All sites were then averaged using a cosine-latitude weighting scheme. Arsenic, chromium, and nickel annual mean concentrations at these sites were all below IMPROVE method detection limits; the estimates of RCE were thus considered upper limits for these pollutants. For chromium VI, the total chromium estimate was multiplied by the average ratio of Chromium VI:Chromium TSP (0.0125) seen in air toxics archive measurements (see **Appendix B**).

Table 4. Remote network concentration estimates. Site locations are defined in Section 2.1.1.

Pollutant	RCE ($\mu g/m^3$)	Remote Network	Year(s)	Location(s)
Chloroform	0.058	AGAGE	2011	Trinidad Head
Methyl chloride (chloromethane)	1.09	AGAGE	2011	Trinidad Head
Chromium VI	1.5E-06	IMPROVE Cr and NATTS Cr VI:Cr ratio	2010-2012	DENA, KALM, PORE, REDW, TRCR, TUXE, HACR
Arsenic	<2.0E-4	IMPROVE	2010-2012	DENA, KALM, PORE, REDW, TRCR, TUXE, HACR
Chromium	<1.2E-4	IMPROVE	2010-2012	DENA, KALM, PORE, REDW, TRCR, TUXE, HACR
Lead	6.6E-04	IMPROVE	2010-2012	DENA, KALM, PORE, REDW, TRCR, TUXE, HACR
Manganese	3.2E-04	IMPROVE	2010-2012	DENA, KALM, PORE, REDW, TRCR, TUXE, HACR
Nickel	<1.0E-4	IMPROVE	2010-2012	DENA, KALM, PORE, REDW, TRCR, TUXE, HACR
Benzene	0.116	NOAA CMDL	2011	KUM, MLO, NWR, BRW, ALT
Carbon tetrachloride	0.547	NOAA CMDL	2011	KUM, MLO, NWR, BRW, ALT
Methyl bromide (bromomethane)	0.0294	NOAA CMDL	2011	KUM, MLO, NWR, BRW, ALT
Methyl chloroform (1,1,1-trichloroethane)	0.06	NOAA CMDL	2011	KUM, MLO, NWR, BRW, ALT
Dichloromethane (methylene chloride)	0.146	NOAA CMDL	2011	KUM, MLO, NWR, BRW, ALT
Tetrachloroethene (perchloroethylene, tetrachloroethylene)	0.0131	NOAA CMDL	2011	KUM, MLO, NWR, BRW, ALT

3.2 Ambient National Network Estimates

The ambient-based method was applied to 36 HAPs. The results of the ambient-based method provide county-level background concentration estimates for counties in the United States and associated territories. Table 5 provides summary statistics for the pollutants to which this method was applied. These summary statistics include the number of counties with ambient measurements between 2010 and 2012, and the minimum, mean, maximum, and 10th, 25th, and 50th percentile background concentrations estimated across these counties. The RCE from the ambient method is the value in the last column of the table; if null, the ambient method was not used. This new RCE is the mean value of the subset of counties below the initial RCE and is unlikely to match any of the summary statistics in the other fields of the table.

In addition to the summary statistics for the background estimates, the table lists the number of counties with "reliable" background estimates that were lower than the RCE from the best of the remote network, literature, or emissions-based methods. If more than three counties were lower than the RCE, we characterized whether these background estimates were within measurement uncertainty (~30%) of the other RCEs from other methods. For carbon tetrachloride, chloromethane, chromium VI, chloroform, and acetonitrile, the background estimates from the ambient method were within this range and were therefore considered consistent with the RCE. In contrast, estimates for the metals of arsenic, cadmium, beryllium, selenium, and cobalt were all lower than the RCEs from other methods by more than 30%. The RCE from the ambient method was considered a more reliable estimate for these pollutants. Because we did not override the literature- or emissions-based estimates if the ambient network estimates were higher, some estimates of remote background concentrations are lower than ambient network estimates would suggest. However, given that the ambient network is primarily in urban areas, it is likely that even the lowest concentrations generated with this method are influenced by local and regional emissions.

Table 5. Summary statistics for the ambient network estimates of county background concentrations. The number of counties with lower RCE estimates indicates that the ambient method generated lower RCE values than those generated from the emissions, literature, or remote network method; if these values were outside of a 30% range of the other RCE value, the new RCE was generated based on the ambient national network method. The "Within 30% of RCE" column was filled only for sites with at least four counties below RCE.

Pollutant	Minimum (µg/m^3)	Pct 10 (µg/m^3)	Pct 25 (µg/m^3)	Pct 50 (µg/m^3)	Mean (µg/m^3)	Number of Counties	No. of Counties with Lower RCE	Within 30% of RCE	New RCE (µg/m^3)
Carbon tetrachloride	7.0E-02	4.4E-01	5.0E-01	5.5E-01	5.1E-01	76	24	Yes	
Beryllium (PM$_{10}$)	1.8E-06	3.5E-06	8.6E-06	1.2E-05	1.1E-05	32	19	No	8.5E-06
Chloromethane (methyl chloride)	8.8E-01	9.7E-01	1.0E+00	1.1E+00	1.1E+00	47	19	Yes	
Cadmium (PM$_{10}$)	1.2E-06	2.6E-05	3.5E-05	4.4E-05	4.9E-05	32	17	No	3.8E-05
Arsenic (PM$_{10}$)	2.5E-05	1.1E-04	1.7E-04	2.0E-04	2.2E-04	35	15	No	1.4E-04
Chromium VI	1.5E-06	1.5E-06	1.5E-06	1.5E-06	5.1E-06	30	13	Yes	
Selenium (PM$_{10}$)	1.1E-04	1.2E-04	1.5E-04	2.4E-04	3.2E-04	18	13	No	2.0E-04
Cobalt (PM$_{10}$)	2.5E-07	2.1E-05	3.8E-05	5.9E-05	1.2E-04	18	11	No	4.1E-05
Chloroform	5.0E-02	5.8E-02	5.8E-02	5.8E-02	6.2E-02	105	7	Yes	
Acetonitrile	1.2E-01	1.7E-01	2.1E-01	2.8E-01	1.1E+00	28	6	Yes	
Lead (PM$_{10}$)	2.5E-04	6.6E-04	6.6E-04	1.0E-03	1.3E-03	46	3	NA	
Antimony (PM$_{10}$)	5.2E-05	6.2E-05	2.5E-04	3.6E-04	6.0E-04	24	2	NA	
Formaldehyde	2.9E-01	7.2E-01	9.1E-01	1.2E+00	1.2E+00	79	2	NA	
Dichloromethane (methylene chloride)	1.4E-01	1.5E-01	1.5E-01	1.7E-01	2.0E-01	112	2	NA	
Nickel (PM$_{10}$)	8.2E-06	1.0E-04	1.0E-04	4.7E-04	6.4E-04	37	2	NA	

Pollutant	Minimum ($\mu g/m^3$)	Pct 10 ($\mu g/m^3$)	Pct 25 ($\mu g/m^3$)	Pct 50 ($\mu g/m^3$)	Mean ($\mu g/m^3$)	Number of Counties	No. of Counties with Lower RCE	Within 30% of RCE	New RCE ($\mu g/m^3$)
Acetaldehyde	1.4E-01	3.9E-01	5.0E-01	7.5E-01	7.5E-01	81	1	NA	
Manganese (PM10)	7.8E-05	5.7E-04	1.1E-03	1.7E-03	2.3E-03	36	1	NA	
Trichloroethene (trichloroethylene)	4.1E-03	4.1E-03	4.1E-03	4.1E-03	4.2E-03	114	1	NA	
1,3-Butadiene	2.0E-03	2.0E-03	2.0E-03	2.0E-03	6.3E-03	121	0	NA	
2,2,4-Trimethylpentane (iso-octane)	9.2E-02	9.2E-02	1.3E-01	1.6E-01	1.8E-01	40	0	NA	
4-Methyl-2-pentanone	4.6E-03	4.6E-03	4.6E-03	4.6E-03	1.8E-02	48	0	NA	
Acrylonitrile	3.6E-04	3.6E-04	3.6E-04	3.6E-04	1.7E-03	40	0	NA	
Benzene	1.1E-01	1.8E-01	2.4E-01	3.2E-01	3.3E-01	126	0	NA	
Carbon disulfide	5.4E-03	5.4E-03	5.4E-03	9.7E-03	4.9E-02	55	0	NA	
Ethylbenzene	1.6E-02	1.6E-02	1.6E-02	4.9E-02	6.4E-02	116	0	NA	
Ethylene dichloride (1,2-dichloroethane)	2.0E-03	2.0E-03	2.0E-03	2.0E-03	1.2E-02	105	0	NA	
Hexane	1.1E-01	1.1E-01	1.4E-01	1.9E-01	2.5E-01	60	0	NA	
Isopropylbenzene (cumene)	3.7E-04	3.7E-04	3.7E-04	3.7E-04	2.4E-03	57	0	NA	
Naphthalene (total tsp &	1.9E-04	6.6E-03	1.0E-02	2.4E-02	2.8E-02	33	0	NA	
1,4-Dichlorobenzene (p-dichlorobenzene)	9.2E-03	9.2E-03	9.2E-03	9.2E-03	1.5E-02	81	0	NA	
Propanal (propionaldehyde)	9.8E-05	9.8E-05	5.3E-02	1.0E-01	1.0E-01	56	0	NA	
Styrene	8.6E-04	8.6E-04	8.6E-04	8.6E-04	8.3E-03	115	0	NA	

Pollutant	Minimum ($\mu g/m^3$)	Pct 10 ($\mu g/m^3$)	Pct 25 ($\mu g/m^3$)	Pct 50 ($\mu g/m^3$)	Mean ($\mu g/m^3$)	Number of Counties	No. of Counties with Lower RCE	Within 30% of RCE	New RCE ($\mu g/m^3$)
Tetrachloroethene (perchloroethylene, tetrachloroethylene)	1.3E-02	1.3E-02	1.3E-02	1.3E-02	2.5E-02	115	0	NA	
Toluene	4.1E-02	1.5E-01	2.6E-01	4.2E-01	4.4E-01	119	0	NA	
Vinyl acetate	3.5E-05	3.5E-05	3.5E-05	2.0E-01	2.9E-01	25	0	NA	
Xylenes	2.3E-02	2.3E-02	6.6E-02	1.9E-01	2.2E-01	114	0	NA	

3.3 Literature-Based Remote Estimates

A literature survey was performed to obtain remote Northern Hemisphere background concentrations that are most likely to represent annual average concentrations at sites not impacted by local or regional emissions of that pollutant. Numerous monitoring studies are available that report short-term or long-term average air toxic pollutant concentrations. We focused on obtaining remote concentrations for the pollutants likely to be of highest risk based on NATA 2005 modeling results. Table 6 outlines the literature survey results. When multiple literature sources were available, the best estimate was selected according to criteria including

- Year of measurement (2011 being the most appropriate)
- Monitoring locations upwind or within the continental U.S. (e.g., Pacific Ocean)
- Satellite-based modeling efforts

Pollutants for which remote concentration estimates were found in the literature are listed in Table 6. This table also provides the location of the measurements, the year(s) of the measurements, and the remote concentration measured.

Table 6. Literature studies that reported pollutant concentrations for air toxics of interest.

Pollutant	RCE (µg/m^3)	Year(s)	Location(s)	First Author, Year of Publication	Citation
1,3-butadiene	0.002	2005	Jungfraujoch, Switzerland	Loov, 2008	DOI: 10.1029/2007JD009751
Acetaldehyde	0.14	2004	Pacific Ocean modeled	Millet, 2010	DOI:10.5194/acp-10-3405-2010
Acrolein	0.016	2005	Jungfraujoch, Switzerland	Loov, 2008	DOI: 10.1029/2007JD009751
Bromoform	0.01	1994-2004	Ocean cruises	Butler, 2007	DOI: 10.1029/2006GB002732
Ethylbenzene	0.016	2005	Jungfraujoch, Switzerland	Legreid, 2008	DOI: 10.1029/2007JD009751
Formaldehyde	0.43	2006	Pacific Ocean satellite, flights, model	Boeke, 2011	DOI: 10.1029/2010JD014870
Toluene	0.041	2005	Jungfraujoch, Switzerland	Loov, 2008	DOI: 10.1029/2007JD009751
Xylenes	0.023	2002	Trinidad Head, California	Millet, 2004	DOI: 10.1029/2003JD004026
Mercury (gaseous)	0.0001	2006	Mercury Deposition network	Butler, 2007	www.arl.noaa.gov/documents/reports/MDN_report.pdf
Acetonitrile	0.17	2004-2008	New Hampshire	Jordan, 2009	DOI: 10.5194/acp-9-4677-2009

Gaseous mercury measurements made in the Mercury Deposition Network are not available in the air toxics monitoring archive and are thus considered "Literature" for the purpose of this report. If data had been in the air toxics monitoring archive, it would have been considered an ambient network pollutant.

3.4 Emissions-Based Estimates

Table 7 provides the emissions, residence times, and estimated remote concentrations for the air toxics included in this study. Measured remote concentration estimates were used for comparison/validation of the method and are also shown in Table 7.

The emissions-based method is not an appropriate method of estimating background concentrations for pollutants with very long residence times in the atmosphere (>350 days) and/or secondary production in the atmosphere (e.g., formaldehyde, acetaldehyde, acrolein), and may be inaccurate for pollutants that are emitted in Asia at very different rates than in the U.S. Long residence time pollutants that last multiple years in the atmosphere include carryover from previous year’s emissions and thus build up over time relative to pollutants with residence times of days to months.

Table 7. Emissions-based method calculated remote background concentrations, 2011 NEI emissions, residence times, and best remote concentration estimates. RCE is an acronym for Remote concentration estimates. Best 2011 RCE estimate is the final background estimate. The table is sorted from highest to lowest Best 2011 RCE.

Name	Residence Time (days)	2011 NEI Emissions v2 (tons per year)	Emissions x Residence Time (tons)	Fraction of C_2Cl_4	Measured Remote Conc 2011 (μg/m^3)	Estimated Remote Conc 2011 (μg/m^3)	Best 2011 RCE estimate (μg/m^3)
Chloromethane (methyl chloride)	365	13,349	4872385	61.07	1.09	0.80	1.09
Carbon tetrachloride	10950	107	1171650	14.68	0.547	0.192	0.55
Formaldehyde	0.13	1371230	178260	2.23	0.43	0.03	0.43
Acetonitrile	365	433	158045	1.98	0.17	0.026	0.170
Acrolein	0.5	52,038	26019	0.33	0.016	0.004	0.0160
Dichloromethane (methylene chloride)	30	8,727	261810	3.28	0.146	0.043	0.146
Acetaldehyde	1	826915	826915	10.36	0.14	0.14	0.14
Benzene	3	279,718	839154	10.52	0.116	0.138	0.12
n-Hexane	2.6	266,516	692942	8.68		0.114	0.11
2,2,4-Trimethylpentane (iso-octane)	4	141,463	565852	7.09		0.093	0.093
Methyl chloroform (1,1,1 trichloroethane)	1825	22,690	41409250	518.99	0.06	6.80	0.06
Chloroform	80	783	62640	0.79	0.058	0.010	0.058
Toluene	0.5	650,831	325416	4.08	0.041	0.053	0.041

Name	Residence Time (days)	2011 NEI Emissions v2 (tons per year)	Emissions x Residence Time (tons)	Fraction of C_2Cl_4	Measured Remote Conc 2011 (µg/m^3)	Estimated Remote Conc 2011 (µg/m^3)	Best 2011 RCE estimate (µg/m^3)
Methyl bromide (bromomethane)	365	6,167	2250955	28.21	0.0294	0.37	0.03
Xylenes	0.2	376,110	75222	0.94	0.023	0.012	0.023
Ethylbenzene	1.7	76,774	130516	1.64	0.016	0.021	0.016
Tetrachloroethene (perchloroethylene, Tetrachloroethylene)	6.5	12,275	79788	1.00	0.0131	0.013	0.013
PAH_880 E5[a]	10	6284	62837	0.79		1.0E-02	1.0E-02
Bromoform	540	34	18349	0.23	0.01	0.0030	1.0E-02
1,4-Dichlorobenzene (p-dichlorobenzene)	31	1,821	56451	0.71		0.009	0.009
Bis(2-ethylhexyl) phthalate	200	175	35000	0.44		0.0057	5.7E-03
Carbon disulfide	7	4,752	33264	0.42		0.0055	5.5E-03
Methyl isobutyl ketone	1	27,977	27977	0.35		0.0046	4.6E-03
Trichloroethene (trichloroethylene)	6	3,374	20244	0.25	0.0041	0.0033	4.1E-03
PAH_176 E4 [a]	10	1680	16803	0.21		2.8E-03	2.8E-03
Ethylene dichloride (1,2-dichloroethane)	42	295	12390	0.16		0.0020	2.0E-03
1,3-Butadiene	0.08	61,576	4926	0.06	0.002	0.0008	2.0E-03
PAH_176 E3 [a]	10	984	9844	0.12		1.6E-03	1.6E-03

Name	Residence Time (days)	2011 NEI Emissions v2 (tons per year)	Emissions x Residence Time (tons)	Fraction of C_2Cl_4	Measured Remote Conc 2011 (µg/m^3)	Estimated Remote Conc 2011 (µg/m^3)	Best 2011 RCE estimate (µg/m^3)
PAH_176 E5 [a]	10	683	6830	0.09		1.1E-03	1.1E-03
Styrene	0.25	21,102	5276	0.07		8.7E-04	8.7E-04
1,1,2,2-Tetrachloroethane	91.3	56	5113	0.06		8.4E-04	8.4E-04
Lead	10	3,105	31050	0.39	6.60E-04	0.0051	6.6E-04
1,3-Dichloropropene	1.25	3,036	3795	0.05		6.2E-04	6.2E-04
Naphthalene	0.25	11,881	2970	0.04		4.9E-04	4.9E-04
1,1,2-Trichloroethane	49	48	2352	0.03		3.9E-04	3.9E-04
Isopropylbenzene (cumene)	2.2	1,035	2277	0.03		3.7E-04	3.7E-04
Ethylene oxide	7	298	2086	0.03		3.4E-04	3.4E-04
Manganese	10	999	9990	0.13	0.000323	0.0016	3.2E-04
Acrylonitrile	5.6	351	1966	0.02		3.2E-04	3.2E-04
1,2-Dichloropropane (propylene dichloride)	30	58	1740	0.02		2.9E-04	2.9E-04
Selenium	10	287	2870	0.04	2.00E-04	4.7E-04	2.0E-04
Ethylene dibromide (1,2-dibromomethane)	50	23	1150	0.01		1.9E-04	1.9E-04
Benzyl chloride	3	291	873	0.01		1.4E-04	1.4E-04
Arsenic	10	126	1260	0.02	1.40E-04	2.1E-04	1.4E-04
Vinyl chloride	2	354	708	0.01		1.2E-04	1.2E-04

Name	Residence Time (days)	2011 NEI Emissions v2 (tons per year)	Emissions x Residence Time (tons)	Fraction of C_2Cl_4	Measured Remote Conc 2011 (µg/m^3)	Estimated Remote Conc 2011 (µg/m^3)	Best 2011 RCE estimate (µg/m^3)
Mercury (gaseous)	365	56.0	20440	0.26	1.00E-04	3.4E-03	1.0E-04
Nickel	10	943	9430	0.12	1.00E-04	0.0015	1.0E-04
Propanal (propionaldehyde)	0.1	5,960	596	0.01		9.8E-05	9.8E-05
Antimony	10	40.0	400	0.01		6.6E-05	6.6E-05
Cobalt	10	57	570	0.01	4.10E-05	9.4E-05	4.1E-05
Cadmium	10	29	290	0.00	3.80E-05	4.8E-05	3.8E-05
Vinyl acetate	0.25	857	214	0.00		3.5E-05	3.5E-05
Beryllium	10	8	79	0.00	8.50E-06	1.3E-05	8.5E-06
PAH_192 E3 [a]	10	3	32	0.00		5.2E-06	5.2E-06
Chromium VI	3	76	229	0.00	1.50E-06	3.8E-05	1.5E-06
1,2-Dibromo-3-chloropropane	36	0.22	8	0.00		1.3E-06	1.3E-06
PAH_114 E1 [a]	10	1	7	0.00		1.1E-06	1.1E-06
PAH_101 E2 [a]	10	0	1	0.00		1.3E-07	1.3E-07
PAH_176 E2 [a]	10	0	1	0.00		9.7E-08	9.7E-08
Hydrazine	0.25	1	0	0.00		5.7E-08	5.7E-08
Benzidine	1	0.15	0	0.00		2.5E-08	2.5E-08

[a] Background concentrations for PAHs are not used because of the uncertainty – the residence time of 10 days is based on a $PM_{2.5}$ residence time and is likely an overestimate.

4. Discussion

4.1 Cancer Risk and Noncancer Hazard

The best 2011 RCEs displayed in Table 7 were used to estimate Northern Hemisphere cancer risk levels from background concentrations. Cancer risk levels were obtained from the EPA Office of Air Quality Planning and Standards (OAQPS) dose-response assessment value and a mutagenicity adjustment factor for cancer risk applied in the HAPEM model in NATA.[10,11] Total remote background cancer risk is approximately 16-in-a-million for all pollutants examined; those shown in **Figure 1** account for about 95% of the total. Of the pollutants listed in Table 7, only four had estimated background cancer risk values above 1-in-a-million. Of these, formaldehyde, carbon tetrachloride, and the polycyclic aromatic hydrocarbon (PAH) group PAH_176 E3 dominated the total cancer risk, as shown in Figure 1. However, it is important to note that acetaldehyde, formaldehyde, and PAH RCEs generated in this work will not be used in NATA 2011. Acetaldehyde and formaldehyde are covered by the CMAQ model, while the PAH RCEs were based on unreliable residence time (10 days, which is the same as $PM_{2.5}$). Thus the actual NATA 2011 risk from background will likely be lower than those estimated here.

Formaldehyde and acetaldehyde are both photochemically produced from precursor volatile organic compounds (VOC) throughout the atmosphere. These pollutants are not transported across the United States; carbonyls are constantly being created and destroyed through atmospheric photo-oxidation processes. The RCE values for these pollutants are representative of the concentrations in areas remote from local and regional VOC emissions. However, since these pollutants are used in CMAQ, and an average secondary production estimate is being applied to non-CMAQ areas (Alaska, Hawaii, Puerto Rico, and the Virgin Islands) it was decided not to apply these background estimates to avoid double counting.

Carbon tetrachloride is a globally distributed pollutant that has been phased out as a result of the Montreal Protocol to reduce chlorofluorocarbons and their impact on the stratospheric ozone layer. Its multi-decade-long atmospheric residence time means that concentrations will decline only slowly over time.

PAH remote concentration estimates are based on emissions from the 2011 NEI and an atmospheric residence time of ten days. There is significant uncertainty in the residence time estimate, as this assumes that the PAHs are in the particulate phase in particles of less than 2.5 micrometer aerodynamic diameter. The emissions-based estimates were applied to 48 individual PAH species; these were summed to the PAH risk group level for use in NATA 2011. Within the PAH_176 E3 group,

---

[10] www2.epa.gov/fera/dose-response-assessment-assessing-health-risks-associated-exposure-hazardous-air-pollutants.
[11] www2.epa.gov/fera/download-human-exposure-model-hem.

which has a total RCE risk of 2.8-in-a-million, methylchrysene is contributing about 77% of the background risk, and benzo[a]pyrene is responsible for the other 23%. For the PAH_880 E5 group, 20 PAHs are included, of which fluoranthene, acenaphthylene, benzo[g,h,i]perylene, and benzo[c]phenanthrene each contribute more than 0.1-in-a-million risk to the total background risk. Due to the uncertainty in the approach, the emissions-based background was not used; instead, a value of 0 was used. This may result in an underestimated risk from transported and background PAHs.

Benzene has a background risk of 0.9-in-a-million. This estimate is based on annual mean measurements at a remote network and is consistent with previous remote concentration estimates. Arsenic has a background risk of about 0.6-in-a-million. This estimate is based on ambient measurements from National Air Toxics Trends Stations (NATTS). Since most of the NATTS locations are urban, this RCE may be skewed high.

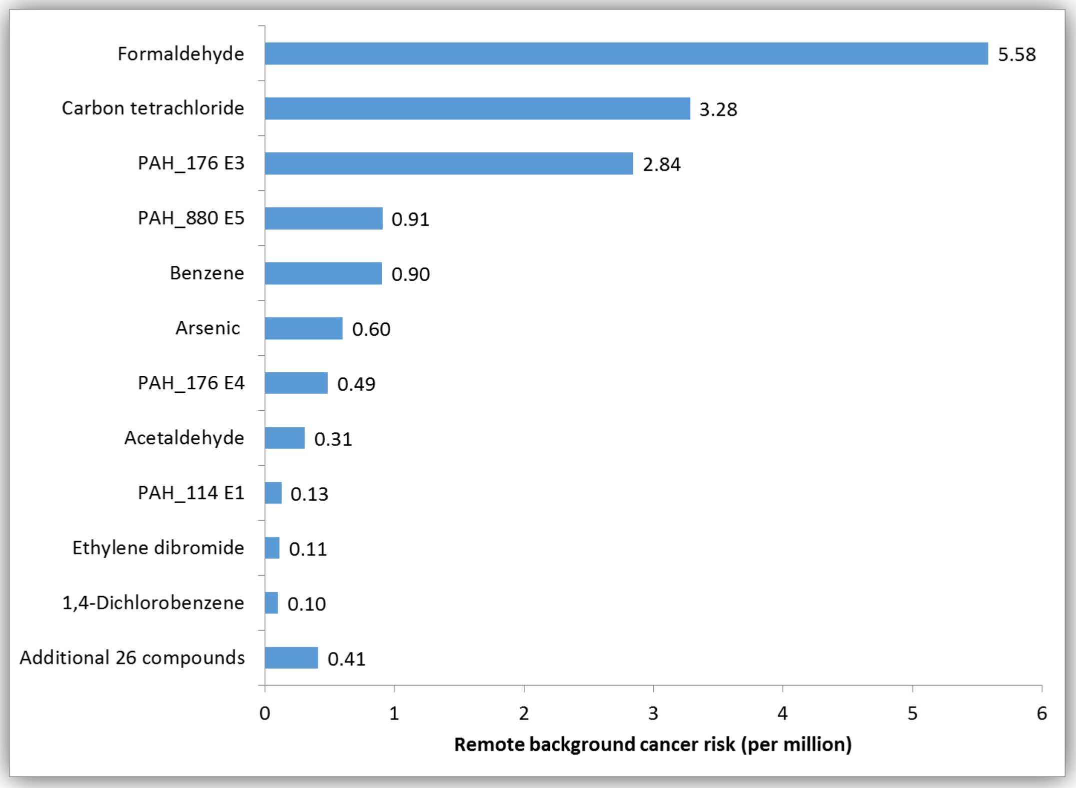


Figure 1. Remote Northern Hemisphere estimates of cancer risk based on concentration estimates from Table 7. PAH categories are consistent with groupings used in NATA 2011. Note that these estimates are not directly applied to final NATA 2011 background estimates.

Noncancer hazard quotients (HQs) were also calculated for the RCEs. Acrolein dominates remote concentration hazards, with a HQ of 0.8. This estimate is based on a literature estimate; emissions-based estimates are far lower but are unable to capture secondary formation of acrolein from 1,3-butadiene photo-oxidation.

Noncancer hazard quotients for all other pollutants were below 0.05. Formaldehyde, acetaldehyde, and chloromethane all have HQs between 0.01 and 0.05.

4.2 Spatial Estimates of Risk

In the initial approach to estimating background concentrations, two methods were used to estimate the spatial gradients in background concentrations that are expected to result from regional transport and emissions of pollutants. In the first approach, ambient-based measurements were used to estimate spatial variability in counties with monitoring stations (see Section 2.1.2). In the second approach, the 2011 NEI v2.0 was used to generate emissions-based gradients in expected background concentrations. This approach is described in detail in **Appendix A**. This emissions-based approach summed emissions from counties within a predetermined buffer distance (<500 km) to estimate the relative impact of nearby emissions on that county for every pollutant. An example of the resulting emissions-based gradient map is shown in **Figure 2**.

The ambient-based approach covers only a very small fraction of U.S. counties due to the limited number of ambient measurement sites (<200 out of ~3200). To extrapolate the results to other counties, we attempted to use the ambient-based county measurements in a multipoint regression with county-based normalized emissions gradients. These two data sets could be used to infer what background concentrations for the counties without measurements should be.

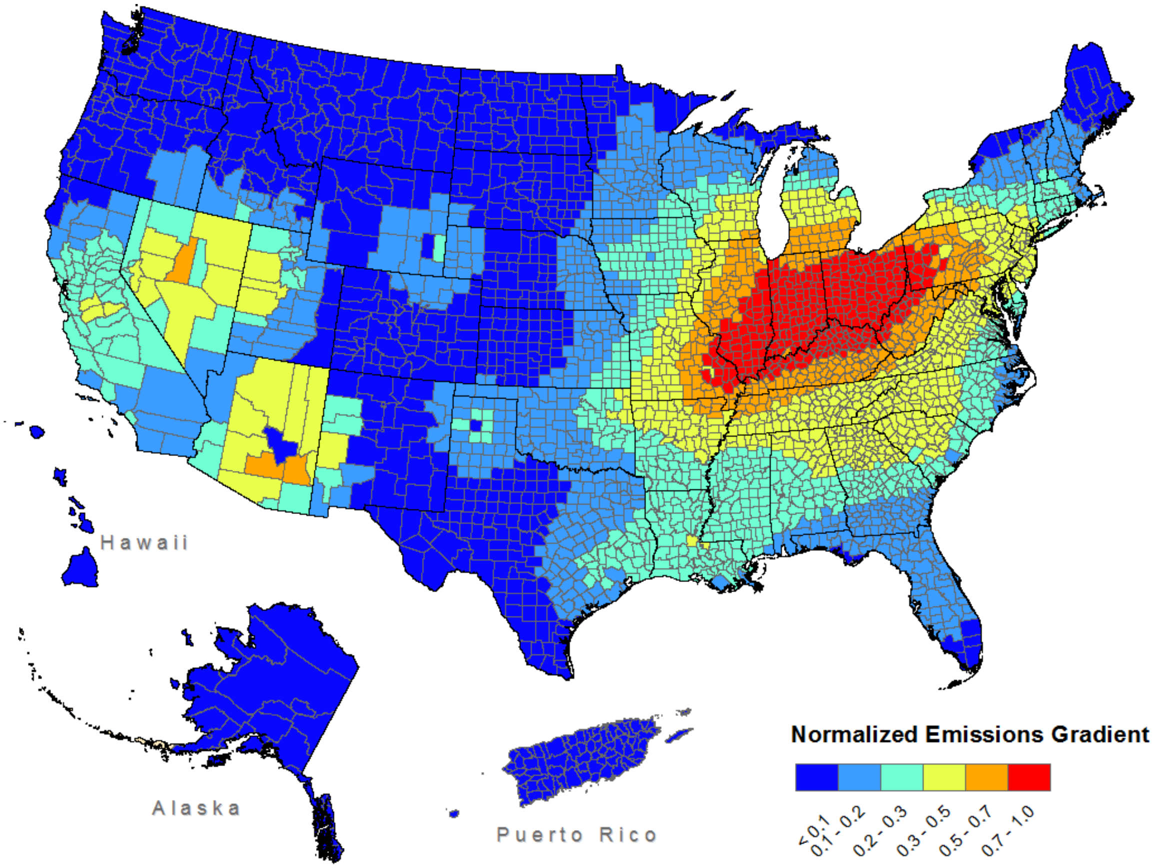


Figure 2. Normalized emissions-based gradient map for arsenic emissions. Counties colored red are expected to have the highest average regional background concentrations; blue counties have the lowest expected regional impacts.

Figure 3 shows regression examples for benzene, arsenic, and carbon tetrachloride, three of the most important cancer risk pollutants. The x-axis (normalized emissions) shows the county emissions-based estimates on a scale of 0 to 1. The y-axis (average concentration background) shows the ambient national network county estimates in units of $\mu g/m^3$. A regression line is fit to each data set, anchored so that the y-intercept must cross at the best remote concentration estimate value from Table 7. In each of these three cases, and in most cases overall, there appears to be no statistically significant relationship between the emissions-based and ambient-based county estimates. For benzene, one of the best measured and characterized air toxics in the United States, Figure 3 shows that low emissions-based method counties have a huge range of background concentrations and the high emissions-based counties do not have high background concentrations. In other words, there is no relationship between the two methods. This poor result could be because of the measurement uncertainty and variability associated with 10th percentile concentrations, or because the emissions-based method does not account for prevailing winds or unrealistic transport distances. Regardless, it is clear that the two methods were incompatible and that predicting spatial variability in regional background concentrations would be based on a method with no statistical backing.

After reviewing these results, the EPA project team decided that a method that attempted only to characterize the remote concentration estimates would be more scientifically and statistically justified than a method that also attempted to predict spatial variability within regional background concentrations. Thus, the approach used in Section 2 was adopted.

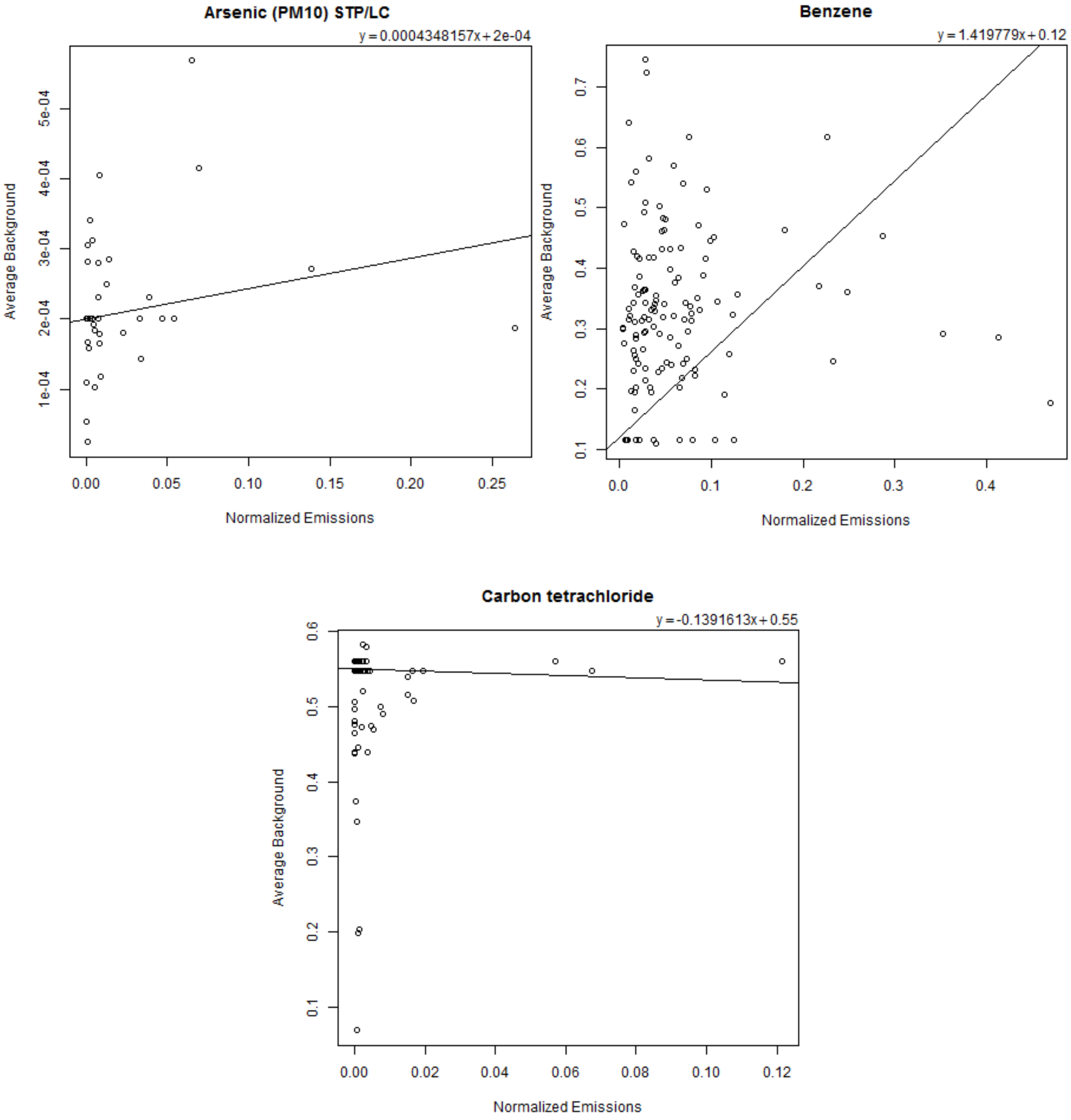


Figure 3. Scatter plots and linear regressions of emissions-based county gradients and ambient network concentrations for arsenic (top left), benzene (top right), and carbon tetrachloride (bottom). The y-intercepts were forced through the best available RCE from the remote network approach.

4.3 Uncertainties

Each of the methods for estimating background concentrations has different levels of uncertainty. The most certain estimates are those from the remote measurement networks, followed by the ambient measurement networks, and then by literature values. The emissions estimates are the most uncertain. Using round numbers, our best estimates for the relative uncertainties are:

- Remote network estimates – highly certain ±25%
- Ambient network estimates – moderate certainty ±50%
- Literature estimates – moderate to low certainty ±50 to 75%
- Emissions estimates – very low certainty ±100%

5. Summary

Remote concentration estimates were determined for 62 pollutants (with PAHs in groups of pollutants). Remote concentration estimates were generated using four distinct technical approaches: remote networks, ambient national network, literature-based, and emissions-based. Each of these approaches was applied to the target pollutants to the extent possible.

Key pollutants with remote concentration estimates exceeding the 1-in-a-million risk level include formaldehyde, carbon tetrachloride, and two groups of PAHs. Remote concentration risk levels for benzene and arsenic were just below 1-in-a-million. Of the 62 pollutants, none had remote concentrations that resulted in a hazard quotient greater than 1.

Spatially varying regional background estimates were generated but did not appear statistically justified, as indicated by regressions between the emissions-based and ambient-based background approaches. Thus, spatially invariant background concentrations were chosen as a more reasonable approach for NATA 2011.

6. References

Bortnick S.M., Coutant B.W., and Biddle B.M. (2003) Estimate background concentrations for the national-scale air toxics assessment. Final technical report prepared for the U.S. Environmental Protection Agency, Research Triangle Park, NC, by Battelle, Columbus, OH, Contract No. 68-D-02-061, Work Assignment 1-03, June.

Hyslop N.P. and White W.H. (2011) Identifying sources of uncertainty from the inter-species covariance of measurement errors. *Environ. Sci. Technol.*, 45(9), 4030-4037, doi: 10.1021/es102605x. Available at http://pubs.acs.org/doi/abs/10.1021/es102605x.

McCarthy M.C., Hafner H.R., and Montzka S.A. (2006) Background concentrations of 18 air toxics for North America. *J. Air and Waste Manag. Assoc.*, 56, 3-11, (STI-903550-2589), January. Available at http://www.awma.org/journal/ShowAbstract.asp?Year=&PaperID=1509.

McCarthy M.C., Rubin J.I., Penfold B.M., and Hafner H.R. (2008) Estimation of background concentrations for NATA 2002. Draft final report prepared for the U.S. Environmental Protection Agency, Research Triangle Park, NC, by Sonoma Technology, Inc., Petaluma, CA, STI-906206.05-3291-DFR, January.

Montzka S.A., Butler J.H., Elkins J.W., Thompson T.M., Clarke A.D., and Lock L.T. (1999) Present and future trends in the atmospheric burden of ozone-depleting halogens. *Nature*, 398, 690-694.

Montzka S.A., Dlugokencky E.J., and Butler J.H. (2011) Non-CO_2 greenhouse gases and climate change. *Nature*, 476(7358), 43-50, doi: 10.1038/nature10322, August 3.

Prinn R.G., Weiss R.F., Fraser P.J., Simmonds P.G., Cunnold D.M., Alyea F.N., O'Doherty S., Salameh P., Miller B.R., Huang J., Wang R.H.J., Hartley D.E., Harth C., Steele L.P., Sturrock G., Midgley P.M., and McCulloch A. (2000) A history of chemically and radiatively important gases in air deduced from ALE/GAGE/AGAGE. *J. Geophys. Res*, 105, 17,751-717,792.

Rosenbaum A.S., Axelrad D.A., Woodruff T.J., Wei Y.H., Ligocki M.P., and Cohen J.P. (1999) National estimates of outdoor air toxics concentrations. *J. Air Waste Manage.*, 49, 1138-1152, (10), Oct.

Woodruff T.J., Axelrad D.A., Caldwell J., Morello-Frosch R., and Rosenbaum A. (1998) Public health implications of 1990 air toxics concentrations across the United States. *Environ. Health Persp.*, 106, 245-251, (5), May.

Appendix A: Emissions-Based Method

The emissions-based method was developed to estimate the spatial variability in regional background concentrations on the basis of spatial differences in county-level emissions. This approach is best applied to pollutants that are emitted directly by a few large sources and that have short residence times in the atmosphere. The emissions-based method consists of four general steps:

1. Import emissions inventory data into a geographic information system (GIS) and create emissions density maps.
2. Apply a spatial weighting scheme for deriving emissions gradients.
3. Normalize the emissions gradients.
4. Convert emissions gradient values to background concentration values.

The emissions-based method uses GIS technology to spatially weight and distribute county-level emissions estimates for each pollutant based on its residence time and air parcel transport potential. These county-level emissions gradient values are then post-processed using lower- and upper-bound anchor points to convert emissions values to background concentrations.

Import Emission Inventory Data into a GIS and Create Emissions Density Maps

The 2011 county-level NEI data were imported into a GIS, and county-level emissions density maps were generated. Because the NEI data consist of a single emissions value for each county by pollutant, it is necessary to spatially distribute the emissions values across county boundaries to account for pollutant transport. To address this, emissions inventory data were spatially weighted and distributed across county boundaries using a distance-residence time weighting scheme for each pollutant. To account for differences in pollutant lifetimes or residence times (i.e., some pollutants remain in the air longer than others), a weighting function was derived and applied within the GIS to create emissions gradients for each pollutant. As an example, Figure A-1 shows the countywide 2011 NEI data for ethylene dibromide. The methodology figures are based on the previous report and have not been altered for this report.

Development of Spatial Weighting Scheme for Deriving Emissions Gradients

The dispersion and dilution assumptions in a Gaussian plume dynamics model lead to concentration dilution of multiple orders of magnitude within a few kilometers. This approach is appropriate for modeling plume movement away from a discrete point source; however, it is less useful for modeling county-level transport. Based on an average wind speed of 3 m/s, air parcel transport is approximately 250 km per day. Consequently, significant transport can occur over two days for pollutants with long residence times. While pollution can be transported farther distances, it is likely that emissions contributions from counties at distances greater than 500 km will be relatively small.

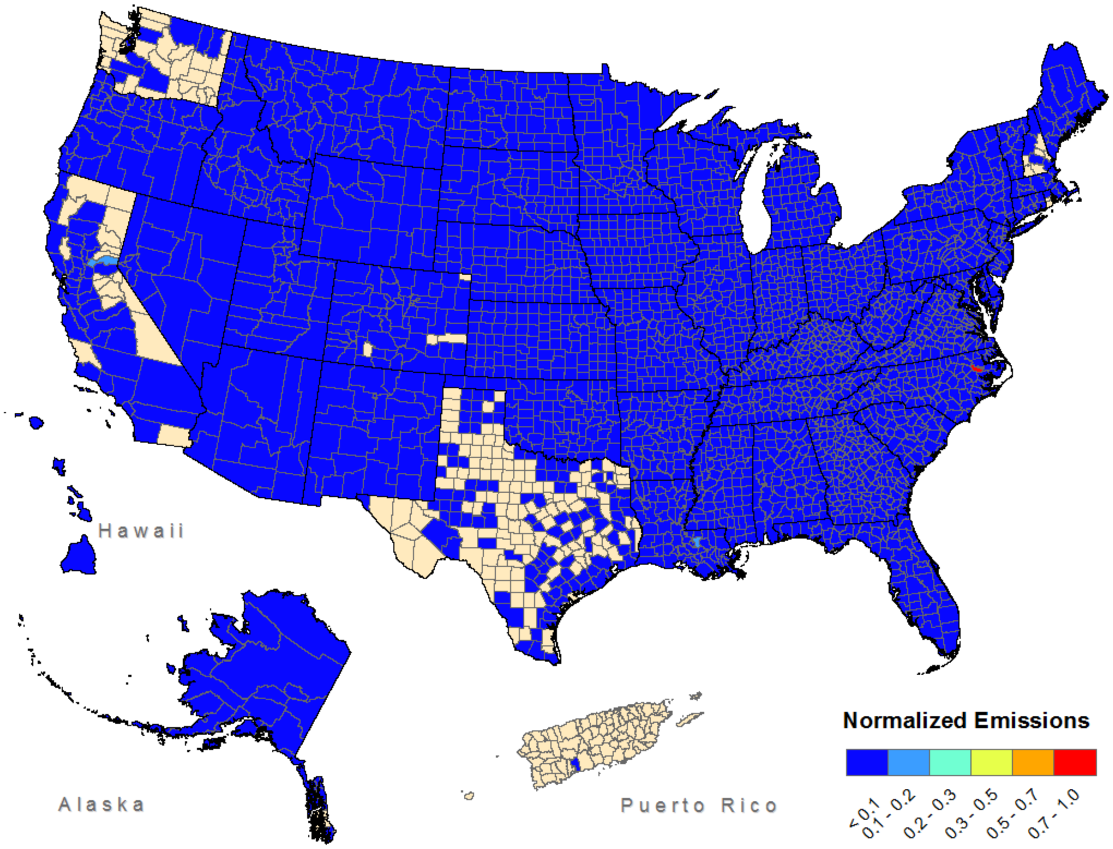


Figure A-1. Countywide ethylene dibromide emissions (tons/year) as reported in the 2011 NEI. Each county is colored according to the magnitude of its total emissions, with tan indicating no reported emissions.

Residence time is another factor contributing to pollutant concentrations over time. Chemical or physical removal competes with dilution if the residence time is on the same order of magnitude as the transport time. If pollutants are removed at rates much slower than they are diluted, they can be treated as inert on the timescale of a few days. In contrast, if pollutants are removed on the timescale of a few hours, the removal processes compete with dilution, and the observed gradient in concentrations is sharper.

For each pollutant of interest, buffer distances based on the residence time and dilution factors were calculated. For the dilution factor of a completely inert pollutant with no deposition, a maximum buffer distance of 500 km was assumed to be the range of influence. While pollution can be transported around the globe, most point source emissions of pollution are fully diluted well within

500 km. This initial 500-km distance was then reduced as a function of the pollutant residence time. Equation A-1 defines the drop-off as a function of distance:

$$B_x = \frac{500}{1.5^{\frac{0.5}{t}}} \tag{A-1}$$

where B_x is buffer distance and t is residence time in days. The exponential equation 1.5^(0.5/t) was empirically selected to provide buffer distances that reflect our expectations. Table A-1 summarizes the buffer distances computed using Equation A-1. Metals in particulate matter were assigned a 10-day residence time, based on estimated residence times of $PM_{2.5}$ in the atmosphere. Because emissions are not broken out by particle size fractions in the NEI, these estimates likely overestimate the range of influence of particulate metals. This approach was chosen because it is more conservative and protective of human health.

Table A-1. Calculated buffer distances for example HAPs of interest.

Pollutants	Buffer Distance (km)
Hydrazine	222
Chromium (VI)	498
Ethylene dichloride (1,2-dichloroethane)	498
Naphthalene	222
1,2-Dichloropropane (propylene dichloride)	496
Ethylene oxide	485
Acrylonitrile	482
Cadmium	490
Beryllium	490
Ethylene dibromide (1,2-dibromomethane)	498
Benzidine	409
Quinoline	499
Bis(2-ethylhexyl)phthalate	500
1,2-Dibromo-3-chloropropane	497
Trichloroethene (trichloroethylene)	483
1,1,2,2-Tetrachloroethane	499
Vinyl chloride	451
Chloroprene	40
Acrolein	333
1,3-Dichloropropene	425

To form a conceptual model of how the buffer distances are applied, consider chloroprene and cadmium. Chloroprene has a relatively short residence time and a resulting buffer distance of 40 km. Assume that the emissions point source for chloroprene is located at the county centroid. As the distance from the county centroid increases, the concentration of chloroprene rapidly decreases due to dilution and chemical reaction. When the distance from the county centroid equals 40 km, it is assumed that the concentration of chloroprene equals zero. Therefore, the contribution of chloroprene from one county to another is likely to be small, because this pollutant has a relatively short residence time. In contrast, cadmium has a much longer residence time and a buffer distance of 490 km. The concentration of cadmium does not reach zero until the distance from the county centroid is 490 km; therefore, the contribution or influence of cadmium from one county to an adjacent one could be relatively high.

For each pollutant, the buffer distance (B_x) was used in **Equation A-2** to estimate the fraction of emissions contribution from a particular county as the distance from the county centroid increases:

$$f_1 = [(B_x - r)/B_x)]^2 \qquad \text{(A-2)}$$

where r is the distance between county centroids, B_x is the distance from the county centroid where the pollutant concentration equals zero, and f_1 is the fraction of emissions contribution from a specific county. The resultant value, f_1, is the fraction of the total emissions of a particular county that are transported to a nearby county.

Calculation Example

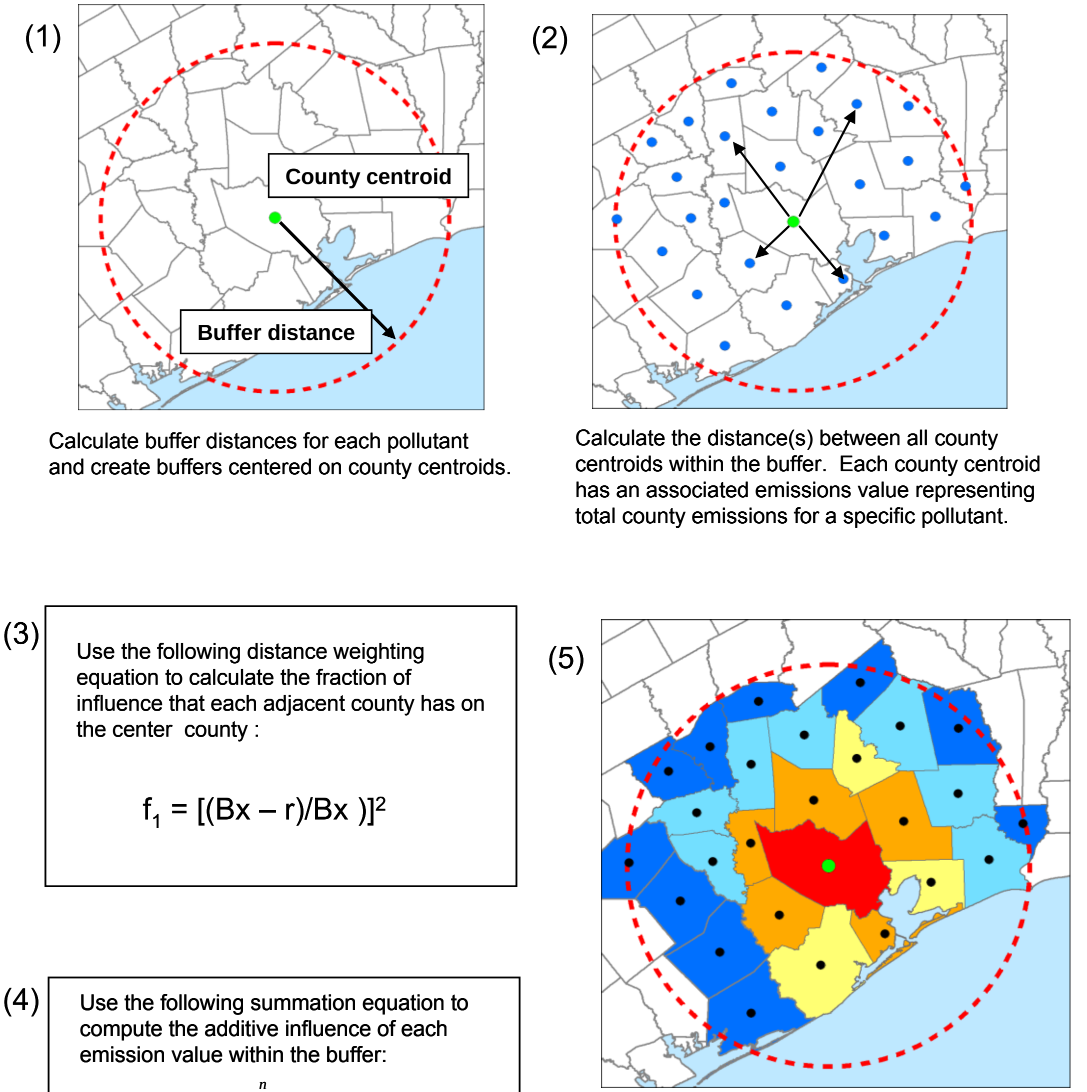
Contribution of chloroprene from County 1, assuming a distance from the centroid of 30 km (r = 30 km):

$f_1 = [(40\text{ km} - 30\text{ km})/40\text{ km})]^2 = (0.25)^2 = 0.063$

Contribution of cadmium from County 1, assuming a distance from the centroid of 30 km (r = 30 km):

$f_1 = [(490\text{ km} - 30\text{ km})/490\text{ km})]^2 = (0.94)^2 = 0.882$

Figure A-2 illustrates the process used to develop and apply the spatial weighting scheme.



(1) Calculate buffer distances for each pollutant and create buffers centered on county centroids.

(2) Calculate the distance(s) between all county centroids within the buffer. Each county centroid has an associated emissions value representing total county emissions for a specific pollutant.

(3) Use the following distance weighting equation to calculate the fraction of influence that each adjacent county has on the center county :

$$f_1 = [(Bx - r)/Bx]^2$$

(4) Use the following summation equation to compute the additive influence of each emission value within the buffer:

$$F_{county} = \frac{\sum_{1}^{n} f_n E_n}{\max(\sum_{1}^{x} f_x E_x)}$$

(5) Spatially distributed emissions values based on the spatial weighting scheme. Note that red indicates areas of high emissions influence and blue represents areas of low influence. Also note that emissions are higher in the center of the buffer zone because as the distance from the county centroid increases the pollutant emissions value decreases as a function of the pollutant buffer distance (listed in Table 5 by pollutant).

Figure A-2. Illustration of the process used to apply the weighting scheme to spatially distribute county-level emissions.

Normalize the Emissions Gradients

Circular buffers centered on a county centroid were created within the GIS. The f_1 values for all counties were calculated within the GIS, and the combined contribution of each county was summed for a given buffer region. Equation A-3 was then used to normalize the emissions contributions from all counties that influence a single county within the buffer zone:

$$\mathrm{F}_{\text{county}} = \frac{\sum_{1}^{n} f_n E_n}{\max(\sum_{1}^{x} f_x E_x)} \tag{A-3}$$

where F_{county} is the county of interest, n is the number of counties with emissions that influence that county, f_n is the fraction emissions value calculated using Equation A-2, E_n is the county emissions value from the 2011 NEI, x is the number of counties that influence the highest emissions county in the country, and *max* indicates the county with the highest emissions in the country for a given pollutant. This calculation is repeated for all counties with reported emissions by pollutant. The weighted emissions values for individual counties were summed and normalized using the county with the maximum emissions contribution (post-calculation). The resulting F_{county} is a unitless value between 0 and 1 representing the lowest and highest transport values in the country, respectively. The normalized F_{county} values were mapped to display the resulting emissions gradient by pollutant. The emissions gradient for each pollutant represents a unitless number corresponding to a range of emissions values. Figure A-3 shows an example of a normalized emissions gradient field for ethylene dibromide.

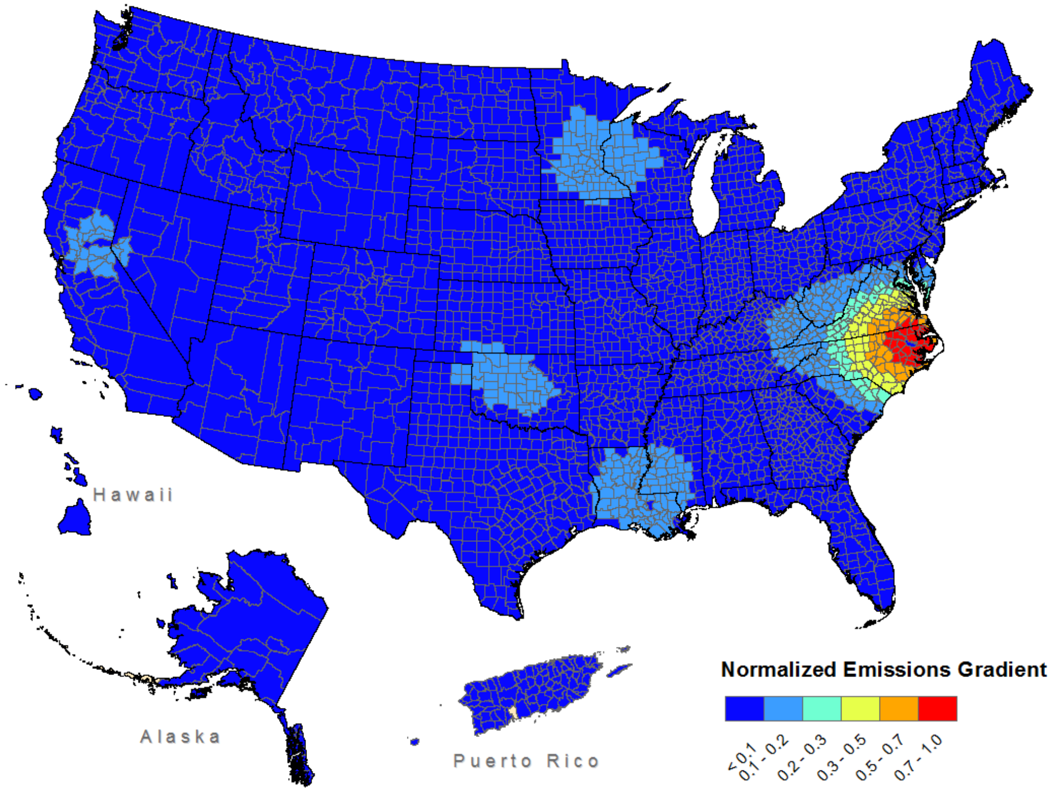


Figure A-3. Final normalized emissions-based gradient map for ethylene dibromide using the 2011 NEI v1.5.

Appendix B: Hexavalent Chromium Ratio Data

The following data were used to determine the ratio of hexavalent chromium to total chromium. The ratio was taken as the average of the mean and the median of the last column of the table, and rounded to the nearest 0.25, so that the value of 1.25% was obtained.

Table 8. Data used to generate the hexavalent chromium to total suspended particulate chromium ratio.

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
010730023	15-Jul-05	0.00204	4.07E-05	2.00%
010730023	21-Jul-05	0.0032	7.23E-05	2.26%
010730023	27-Jul-05	0.00276	9.97E-05	3.61%
010730023	08-Aug-05	0.00365	7.93E-05	2.17%
010730023	20-Aug-05	0.00257	1.55E-05	0.60%
010730023	13-Sep-05	0.00327	2.99E-05	0.91%
010730023	07-Oct-05	0.00174	1.63E-05	0.94%
010730023	19-Oct-05	0.00576	8.17E-05	1.42%
010730023	12-Nov-05	0.00463	5.12E-05	1.11%
010730023	06-Dec-05	0.00269	1.81E-05	0.67%
010730023	30-Dec-05	0.00436	6.24E-05	1.43%
010730023	11-Jan-06	0.00489	5.2E-05	1.06%
010730023	23-Jan-06	0.00407	4.855E-05	1.19%
010730023	16-Feb-06	0.00456	3.04E-05	0.67%
010730023	28-Feb-06	0.00528	4.04E-05	0.77%
010730023	24-Mar-06	0.00452	1.59E-05	0.35%
010730023	05-Apr-06	0.0053	0.000192	3.62%
010730023	29-Apr-06	0.00968	0.000166	1.71%
010730023	11-May-06	0.00456	2.06E-05	0.45%
010730023	23-May-06	0.00485	0.000125	2.58%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
010730023	04-Jun-06	0.00432	2.5E-05	0.58%
010730023	16-Jun-06	0.00685	9.3E-05	1.36%
010730023	28-Jun-06	0.00369	4.58E-05	1.24%
010730028	15-Jul-05	0.00336	2.96E-05	0.88%
010730028	21-Jul-05	0.00417	8.13E-05	1.95%
010730028	27-Jul-05	0.00502	3.8E-05	0.76%
010730028	08-Aug-05	0.00342	6.18E-05	1.81%
010730028	20-Aug-05	0.00504	3.47E-05	0.69%
010730028	13-Sep-05	0.00426	7.61E-05	1.79%
010730028	07-Oct-05	0.00402	3.48E-05	0.87%
010730028	12-Nov-05	0.00294	6.86E-05	2.33%
010730028	24-Nov-05	0.0027	1.76E-05	0.65%
010730028	06-Dec-05	0.00521	5.57E-05	1.07%
010730028	18-Dec-05	0.00324	4.14E-05	1.28%
010730028	11-Jan-06	0.00568	6.38E-05	1.12%
010730028	04-Feb-06	0.00454	4.6E-06	0.10%
010730028	16-Feb-06	0.00501	3.75E-05	0.75%
010730028	28-Feb-06	0.00764	6.41E-05	0.84%
010730028	24-Mar-06	0.00586	2.4E-05	0.41%
010730028	05-Apr-06	0.00778	8.71E-05	1.12%
010730028	17-Apr-06	0.00654	5.82E-05	0.89%
010730028	29-Apr-06	0.00569	2.15E-05	0.38%
010730028	11-May-06	0.00594	2.61E-05	0.44%
010730028	23-May-06	0.00577	8.88E-05	1.54%
010730028	04-Jun-06	0.00552	2.31E-05	0.42%
010730028	16-Jun-06	0.00532	5.96E-05	1.12%
010730028	28-Jun-06	0.00637	0.0002	3.14%
010731009	21-Jul-05	0.00144	2.56E-05	1.78%
010731009	08-Aug-05	0.00125	2.26E-05	1.81%
010731009	20-Aug-05	0.00143	1.92E-05	1.34%
010731009	25-Sep-05	0.00128	2.43E-05	1.90%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
010731009	12-Nov-05	0.00144	1.02E-05	0.71%
010731009	06-Dec-05	0.00165	3.9E-06	0.24%
010731009	30-Dec-05	0.00182	3.7E-06	0.20%
010731009	23-Jan-06	0.00311	5.5E-06	0.18%
010731009	16-Feb-06	0.00346	2.54E-05	0.73%
010731009	05-Apr-06	0.00337	1.56E-05	0.46%
010731009	17-Apr-06	0.00284	1.05E-05	0.37%
010731009	29-Apr-06	0.00425	2.57E-05	0.60%
010731009	11-May-06	0.00365	1.26E-05	0.35%
010731009	04-Jun-06	0.00383	8.8E-06	0.23%
010731009	16-Jun-06	0.00315	5.9E-05	1.87%
010731009	28-Jun-06	0.00269	3.18E-05	1.18%
010736004	21-Jul-05	0.00513	4.5E-05	0.88%
010736004	27-Jul-05	0.00432	9.16E-05	2.12%
010736004	08-Aug-05	0.00199	4.37E-05	2.20%
010736004	20-Aug-05	0.00297	3.29E-05	1.11%
010736004	13-Sep-05	0.00356	4.73E-05	1.33%
010736004	25-Sep-05	0.0025	2.86E-05	1.14%
010736004	12-Nov-05	0.00334	4.25E-05	1.27%
010736004	06-Dec-05	0.00302	0.000104	3.44%
010736004	30-Dec-05	0.00434	3.43E-05	0.79%
010736004	23-Jan-06	0.00247	1.7E-06	0.07%
010736004	16-Feb-06	0.00531	2.52E-05	0.47%
010736004	28-Feb-06	0.0074	6.73E-05	0.91%
010736004	24-Mar-06	0.0109	5.79E-05	0.53%
010736004	05-Apr-06	0.00475	0.000125	2.63%
010736004	17-Apr-06	0.00567	5.2E-05	0.92%
010736004	29-Apr-06	0.00573	0.000198	3.46%
010736004	11-May-06	0.00446	4.185E-05	0.94%
010736004	23-May-06	0.00533	0.000123	2.31%
010736004	04-Jun-06	0.00998	1.58E-05	0.16%

AQS Site Code	Sample Date	Chromium TSP (µg/m³)	Chromium VI (µg/m³)	Chromium VI:Chromium TSP (percentage)
010736004	16-Jun-06	0.00551	0.000107	1.94%
010736004	28-Jun-06	0.0133	0.000213	1.60%
060371103	12-Jan-07	0.0015	0.000118	7.87%
060371103	01-Mar-07	0.0042	0.000139	3.31%
060371103	25-Mar-07	0.0038	0.000129	3.39%
060371103	12-May-07	0.0036	0.000137	3.81%
060371103	24-May-07	0.0041	0.000066	1.61%
060371103	05-Jun-07	0.0036	0.000108	3.00%
060371103	17-Jun-07	0.0039	0.000032	0.82%
060371103	11-Jul-07	0.0039	0.000057	1.46%
060371103	23-Jul-07	0.0049	0.000165	3.37%
060371103	04-Aug-07	0.0036	0.000134	3.72%
060371103	16-Aug-07	0.0057	0.000173	3.04%
060371103	28-Aug-07	0.0057	0.0003235	5.68%
060371103	09-Sep-07	0.0033	0.000043	1.30%
060371103	21-Sep-07	0.0045	0.000194	4.31%
060371103	03-Oct-07	0.0051	0.000103	2.02%
060371103	15-Oct-07	0.0034	0.000198	5.82%
060371103	08-Nov-07	0.0038	0.000074	1.95%
060371103	20-Nov-07	0.0047	0.000029	0.62%
060371103	02-Dec-07	0.0036	0.000087	2.42%
060371103	14-Dec-07	0.0032	0.000196	6.13%
060371103	26-Dec-07	0.0036	0.00013	3.61%
060371103	01-Jan-09	0.0048	0.000106	2.21%
060371103	13-Jan-09	0.0079	0.000297	3.76%
060371103	18-Feb-09	0.0035	0.000097	2.77%
060371103	02-Mar-09	0.0054	0.000188	3.48%
060371103	26-Mar-09	0.0041	0.000063	1.54%
060371103	07-Apr-09	0.0034	0.00008	2.35%
060371103	19-Apr-09	0.0036	0.000094	2.61%
060371103	01-May-09	0.006	0.000057	0.95%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
060371103	13-May-09	0.0051	0.000059	1.16%
060371103	18-Jun-09	0.0044	0.00004	0.91%
060371103	30-Jun-09	0.0033	0.000139	4.21%
060371103	12-Jul-09	0.0033	0.000089	2.70%
060371103	24-Jul-09	0.0056	0.000096	1.71%
060371103	22-Sep-09	0.0059	0.000125	2.12%
060371103	04-Oct-09	0.0037	0.000074	2.00%
060371103	16-Oct-09	0.0077	0.000122	1.58%
060371103	09-Nov-09	0.0059	0.000108	1.83%
060371103	03-Dec-09	0.0055	0.000157	2.85%
060371103	15-Dec-09	0.0034	0.000194	5.71%
060371103	21-Dec-09	0.0032	0.000154	4.81%
060371103	14-Jan-10	0.0056	0.00018	3.21%
060371103	20-Jan-10	0.0015	0.00009	6.00%
060371103	01-Feb-10	0.0052	0.00003	0.58%
060371103	25-Feb-10	0.0034	0.0001	2.94%
060371103	09-Mar-10	0.0015	0.00011	7.33%
060371103	21-Mar-10	0.0046	0.00007	1.52%
060371103	02-Apr-10	0.0037	0.00007	1.89%
060371103	14-Apr-10	0.0038	0.00007	1.84%
060371103	26-Apr-10	0.0032	0.00005	1.56%
060371103	08-May-10	0.0044	0.00005	1.14%
060371103	20-May-10	0.0038	0.00011	2.89%
060371103	01-Jun-10	0.0045	0.00011	2.44%
060371103	13-Jun-10	0.0041	0.00002	0.49%
060371103	25-Jun-10	0.0049	0.00004	0.82%
060371103	07-Jul-10	0.044	0.0001	0.23%
060371103	19-Jul-10	0.0043	0.00011	2.56%
060658001	01-Mar-07	0.0067	0.000145	2.16%
060658001	25-Mar-07	0.0053	0.000185	3.49%
060658001	18-Apr-07	0.0047	0.000132	2.81%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
060658001	16-Aug-07	0.0067	0.00017	2.54%
060658001	22-Aug-07	0.0057	0.000241	4.23%
060658001	28-Aug-07	0.0049	0.0004515	9.21%
060658001	09-Sep-07	0.0038	0.000053	1.39%
060658001	03-Oct-07	0.0069	0.001039	15.06%
060658001	15-Oct-07	0.0015	0.000052	3.47%
060658001	26-Nov-07	0.0065	0.001307	20.11%
060658001	02-Dec-07	0.0015	0.000385	25.67%
060658001	14-Dec-07	0.0015	0.000988	65.87%
060658001	26-Dec-07	0.0033	0.000688	20.85%
060658001	01-Jan-09	0.0043	0.000405	9.42%
060658001	13-Jan-09	0.0015	0.000043	2.87%
060658001	25-Jan-09	0.0015	0.000047	3.13%
060658001	18-Feb-09	0.0034	0.000173	5.09%
060658001	02-Mar-09	0.0047	0.00039	8.30%
060658001	08-Mar-09	0.0015	0.000066	4.40%
060658001	14-Mar-09	0.0031	0.000037	1.19%
060658001	26-Mar-09	0.0039	0.000086	2.21%
060658001	07-Apr-09	0.0049	0.00012	2.45%
060658001	19-Apr-09	0.0051	0.00021	4.12%
060658001	01-May-09	0.0059	0.000066	1.12%
060658001	13-May-09	0.0049	0.000037	0.76%
060658001	25-May-09	0.0042	0.000038	0.90%
060658001	06-Jun-09	0.0035	0.000088	2.51%
060658001	18-Jun-09	0.0045	0.000076	1.69%
060658001	30-Jun-09	0.0055	0.000115	2.09%
060658001	12-Jul-09	0.0036	0.000047	1.31%
060658001	24-Jul-09	0.0053	0.0000785	1.48%
060658001	05-Aug-09	0.0015	0.000223	14.87%
060658001	17-Aug-09	0.0015	0.000046	3.07%
060658001	29-Aug-09	0.0015	0.000163	10.87%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
060658001	10-Sep-09	0.0052	0.00016	3.08%
060658001	22-Sep-09	0.0034	0.000253	7.44%
060658001	04-Oct-09	0.0035	0.000054	1.54%
060658001	16-Oct-09	0.0062	0.000394	6.35%
060658001	28-Oct-09	0.0015	0.000034	2.27%
060658001	09-Nov-09	0.0056	0.000225	4.02%
060658001	21-Nov-09	0.0042	0.000087	2.07%
060658001	03-Dec-09	0.0041	0.000128	3.12%
060658001	15-Dec-09	0.0037	0.000147	3.97%
060658001	27-Dec-09	0.0032	0.000297	9.28%
060658001	08-Jan-10	0.0036	0.00026	7.22%
060658001	20-Jan-10	0.0015	0.00011	7.33%
060658001	01-Feb-10	0.0045	0.0001	2.22%
060658001	13-Feb-10	0.0035	0.0001	2.86%
060658001	25-Feb-10	0.0042	0.00012	2.86%
060658001	09-Mar-10	0.0035	0.00008	2.29%
060658001	21-Mar-10	0.0035	0.00006	1.71%
060658001	02-Apr-10	0.0049	0.00008	1.63%
060658001	14-Apr-10	0.0044	0.00006	1.36%
060658001	26-Apr-10	0.0053	0.00006	1.13%
060658001	08-May-10	0.0057	0.00004	0.70%
060658001	20-May-10	0.0051	0.00009	1.76%
060658001	01-Jun-10	0.0048	0.00004	0.83%
060658001	13-Jun-10	0.0037	0.00006	1.62%
060658001	25-Jun-10	0.0076	0.00015	1.97%
060658001	07-Jul-10	0.0046	0.00007	1.52%
060658001	19-Jul-10	0.0043	0.00007	1.63%
060658001	31-Jul-10	0.0047	0.00006	1.28%
060658001	05-Sep-10	0.0044	0.00004	0.91%
060658001	17-Sep-10	0.0034	0.000035	1.03%
060658001	29-Sep-10	0.0053	0.00005	0.94%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
060658001	10-Dec-10	0.0037	0.00003	0.81%
060658001	22-Dec-10	0.0015	0.00003	2.00%
170314201	28-Jan-05	0.00348	0.0000282	0.81%
170314201	27-Feb-05	0.00281	0.000112	3.99%
170314201	29-Mar-05	0.00341	0.000029	0.85%
170314201	10-Apr-05	0.00254	0.0000227	0.89%
170314201	16-Apr-05	0.00247	0.0000601	2.43%
170314201	22-Apr-05	0.00171	0.0000504	2.95%
170314201	10-May-05	0.00182	0.0000626	3.44%
170314201	16-May-05	0.00159	0.0000423	2.66%
170314201	22-May-05	0.00024	0.000009	3.75%
170314201	15-Jun-05	0.00219	0.0000558	2.55%
170314201	27-Jun-05	0.0033	0.0000507	1.54%
170314201	03-Jul-05	0.00307	0.0000829	2.70%
170314201	15-Jul-05	0.00126	0.0000346	2.75%
170314201	21-Jul-05	0.00127	0.0000456	3.59%
170314201	27-Jul-05	0.0011	0.0000145	1.32%
170314201	02-Aug-05	0.00186	0.0000315	1.69%
170314201	08-Aug-05	0.00218	0.0000312	1.43%
170314201	14-Aug-05	0.00289	0.0000225	0.78%
170314201	20-Aug-05	0.00263	0.0000651	2.48%
170314201	26-Aug-05	0.0043	0.0000229	0.53%
170314201	07-Sep-05	0.00379	0.0000273	0.72%
170314201	25-Sep-05	0.00367	0.000034	0.93%
170314201	11-Jan-06	0.003	0.0000402	1.34%
170314201	17-Jan-06	0.004	0.0000216	0.54%
170314201	23-Jan-06	0.003	0.00002325	0.77%
170314201	17-May-06	0.004	0.0000522	1.30%
170314201	07-Nov-06	0.004	0.0000495	1.24%
170314201	19-Dec-06	0.004	0.0000397	0.99%
170314201	25-Mar-07	0.003	0.0000055	0.18%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
170314201	12-Apr-07	0.003	0.0000046	0.15%
170314201	24-Apr-07	0.003	0.0000088	0.29%
170314201	18-May-07	0.003	0.0000442	1.47%
170314201	24-May-07	0.003	0.0000494	1.65%
170314201	11-Jun-07	0.003	0.0000153	0.51%
170314201	05-Jul-07	0.007	0.000307	4.39%
170314201	17-Jul-07	0.003	0.0000122	0.41%
170314201	04-Aug-07	0.003	0.0000168	0.56%
170314201	03-Sep-07	0.002	0.0000097	0.49%
170314201	09-Sep-07	0.002	0.0000234	1.17%
170314201	15-Sep-07	0.002	0.00001675	0.84%
170314201	15-Oct-07	0.003	0.0000285	0.95%
170314201	20-Dec-07	0.005	0.0000272	0.54%
170314201	26-Dec-07	0.004	0.0000102	0.26%
170314201	13-Jan-08	0.003	0.0000079	0.26%
170314201	06-Apr-08	0.004	0.0000226	0.57%
170314201	18-Apr-08	0.003	0.0000193	0.64%
170314201	24-Apr-08	0.005	0.0000558	1.12%
170314201	30-Apr-08	0.004	0.0000204	0.51%
170314201	06-May-08	0.005	0.0000476	0.95%
170314201	30-May-08	0.003	0.0000356	1.19%
170314201	05-Jun-08	0.004	0.0000159	0.40%
170314201	29-Jul-08	0.003	0.0000123	0.41%
170314201	28-Aug-08	0.006	0.0000248	0.41%
170314201	21-Sep-08	0.006	0.0000292	0.49%
170314201	27-Sep-08	0.004	0.0000174	0.44%
170314201	08-Dec-08	0.003	0.0000302	1.01%
170314201	14-Mar-09	0.003	0.0000243	0.81%
170314201	18-Jun-09	0.003	0.0000337	1.12%
170314201	24-Jun-09	0.003	0.0000629	2.10%
170314201	30-Jul-09	0.003	0.0000343	1.14%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
261630033	10-Jan-05	0.007575	0.000036	0.48%
261630033	28-Jan-05	0.006205	0.0000055	0.09%
261630033	03-Feb-05	0.0110876	0.0000309	0.28%
261630033	15-Feb-05	0.0115975	0.0000752	0.65%
261630033	21-Feb-05	0.0066257	0.0000165	0.25%
261630033	11-Mar-05	0.008995	0.0000334	0.37%
261630033	23-Mar-05	0.03338	0.0000805	0.24%
261630033	29-Mar-05	0.015265	0.0000742	0.49%
261630033	04-Apr-05	0.009915	0.0000587	0.59%
261630033	10-Apr-05	0.00548	0.0000464	0.85%
261630033	16-Apr-05	0.0073	0.000146	2.00%
261630033	22-Apr-05	0.006525	0.000102	1.56%
261630033	04-May-05	0.007195	0.0000714	0.99%
261630033	10-May-05	0.014895	0.000126	0.85%
261630033	22-May-05	0.00567	0.0000854	1.51%
261630033	28-May-05	0.005535	0.0000792	1.43%
261630033	03-Jun-05	0.00683	0.0000659	0.96%
261630033	21-Jun-05	0.008405	0.000136	1.62%
261630033	27-Jun-05	0.0155	0.0000902	0.58%
261630033	03-Jul-05	0.034815001	0.0000891	0.26%
261630033	15-Jul-05	0.005995	0.0000418	0.70%
261630033	21-Jul-05	0.00559	0.0000739	1.32%
261630033	27-Jul-05	0.004675	0.0000265	0.57%
261630033	02-Aug-05	0.00579	0.0000514	0.89%
261630033	08-Aug-05	0.008825	0.0000889	1.01%
261630033	26-Aug-05	0.00842	0.0000665	0.79%
261630033	07-Sep-05	0.009725	0.000105	1.08%
261630033	13-Sep-05	0.01208	0.0000619	0.51%
261630033	19-Sep-05	0.00866	0.0000616	0.71%
261630033	25-Sep-05	0.0069	0.0000776	1.12%
261630033	01-Oct-05	0.008396	0.0000861	1.03%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
261630033	07-Oct-05	0.005262	0.0000545	1.04%
261630033	13-Oct-05	0.0072485	0.0000787	1.09%
261630033	19-Oct-05	0.0096345	0.000068	0.71%
261630033	31-Oct-05	0.0108525	0.0000604	0.56%
261630033	06-Nov-05	0.008167	0.0000581	0.71%
261630033	12-Nov-05	0.0070885	0.0000765	1.08%
261630033	18-Nov-05	0.012625	0.0000271	0.21%
261630033	24-Nov-05	0.005605	0.00002215	0.40%
261630033	30-Nov-05	0.0062765	0.0000268	0.43%
261630033	06-Dec-05	0.005948	0.0000338	0.57%
261630033	12-Dec-05	0.005207	0.0000387	0.74%
261630033	18-Dec-05	0.004663	0.000058	1.24%
261630033	24-Dec-05	0.005546	0.0000882	1.59%
261630033	05-Jan-06	0.0031204	0.0000604	1.94%
261630033	11-Jan-06	0.0067323	0.0000415	0.62%
261630033	17-Jan-06	0.0050784	0.00022	4.33%
261630033	23-Jan-06	0.0100253	0.0000197	0.20%
261630033	29-Jan-06	0.0050225	0.0000559	1.11%
261630033	04-Feb-06	0.0031199	0.000064	2.05%
261630033	10-Feb-06	0.0062883	0.0000482	0.77%
261630033	22-Feb-06	0.0109229	0.0000704	0.64%
261630033	28-Feb-06	0.0037052	0.0000242	0.65%
261630033	06-Mar-06	0.0042686	0.0000236	0.55%
261630033	18-Mar-06	0.0057351	0.00003025	0.53%
261630033	24-Mar-06	0.0048841	0.000232	4.75%
261630033	30-Mar-06	0.0086551	0.0000697	0.81%
261630033	05-Apr-06	0.0050755	0.0000344	0.68%
261630033	11-Apr-06	0.01128215	0.000105	0.93%
261630033	17-Apr-06	0.00471765	0.0000189	0.40%
261630033	23-Apr-06	0.0025276	0.0000578	2.29%
261630033	29-Apr-06	0.0030684	0.0000152	0.50%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
261630033	05-May-06	0.0066617	0.0000333	0.50%
261630033	11-May-06	0.0048568	0.0000288	0.59%
261630033	17-May-06	0.0057822	0.0000336	0.58%
261630033	23-May-06	0.0064621	0.0000777	1.20%
261630033	29-May-06	0.00468725	0.0000388	0.83%
261630033	04-Jun-06	0.0029205	0.0000259	0.89%
261630033	10-Jun-06	0.0033129	0.0000146	0.44%
261630033	16-Jun-06	0.0138529	0.000116	0.84%
261630033	22-Jun-06	0.004676	0.0000687	1.47%
261630033	28-Jun-06	0.0070246	0.0000789	1.12%
261630033	04-Jul-06	0.00638965	0.000496	7.76%
261630033	10-Jul-06	0.0052803	0.0000636	1.20%
261630033	16-Jul-06	0.00700115	0.00005615	0.80%
261630033	22-Jul-06	0.00404065	0.0000236	0.58%
261630033	28-Jul-06	0.00578725	0.0000518	0.90%
261630033	03-Aug-06	0.0042248	0.00012	2.84%
261630033	09-Aug-06	0.00894825	0.00012	1.34%
261630033	21-Aug-06	0.0084533	0.0000761	0.90%
261630033	27-Aug-06	0.00342715	0.0000333	0.97%
261630033	02-Sep-06	0.00376825	0.000103	2.73%
261630033	08-Sep-06	0.0093528	0.0000609	0.65%
261630033	14-Sep-06	0.00366095	0.0000349	0.95%
261630033	20-Sep-06	0.00569815	0.0000396	0.69%
261630033	02-Oct-06	0.0084389	0.000254	3.01%
261630033	08-Oct-06	0.00659315	0.000108	1.64%
261630033	20-Oct-06	0.0059855	0.00006525	1.09%
261630033	26-Oct-06	0.0103236	0.00016	1.55%
261630033	07-Nov-06	0.00494715	0.00011515	2.33%
261630033	13-Nov-06	0.00494715	0.00001565	0.32%
261630033	25-Nov-06	0.00971735	0.0000556	0.57%
261630033	01-Dec-06	0.00413425	0.0000205	0.50%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
261630033	13-Dec-06	0.0070503	0.00002335	0.33%
261630033	19-Dec-06	0.0056269	0.0000125	0.22%
261630033	25-Dec-06	0.0031296	0.00003685	1.18%
261630033	31-Dec-06	0.00572405	0.0000235	0.41%
261630033	06-Jan-07	0.00262085	0.0000043	0.16%
261630033	12-Jan-07	0.0082493	0.0000532	0.64%
261630033	18-Jan-07	0.00732745	0.0000319	0.44%
261630033	24-Jan-07	0.0045273	0.00001565	0.35%
261630033	30-Jan-07	0.0049467	0.0000105	0.21%
261630033	11-Feb-07	0.0080771	0.0000173	0.21%
261630033	17-Feb-07	0.0035706	0.0000193	0.54%
261630033	23-Feb-07	0.0039646	0.0000105	0.26%
261630033	01-Mar-07	0.0022323	0.0000151	0.68%
261630033	07-Mar-07	0.00509755	0.0000158	0.31%
261630033	13-Mar-07	0.0122883	0.000133	1.08%
261630033	19-Mar-07	0.00958605	0.00002665	0.28%
261630033	25-Mar-07	0.0036816	0.0000263	0.71%
261630033	31-Mar-07	0.00320865	0.0000182	0.57%
261630033	06-Apr-07	0.0037542	0.0000118	0.31%
261630033	12-Apr-07	0.0068328	0.0000315	0.46%
261630033	18-Apr-07	0.00597985	0.0000196	0.33%
261630033	24-Apr-07	0.0082121	0.0000511	0.62%
261630033	30-Apr-07	0.0050975	0.0000188	0.37%
261630033	06-May-07	0.0032916	0.0000071	0.22%
261630033	12-May-07	0.00392465	0.000012	0.31%
261630033	18-May-07	0.0061486	0.0000356	0.58%
261630033	24-May-07	0.0087943	0.0000929	1.06%
261630033	30-May-07	0.01143405	0.000113	0.99%
261630033	05-Jun-07	0.0031619	0.0000174	0.55%
261630033	11-Jun-07	0.0075505	0.000037	0.49%
261630033	17-Jun-07	0.0037399	0.0000123	0.33%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
261630033	23-Jun-07	0.00633665	0.0000506	0.80%
261630033	29-Jun-07	0.0050638	0.000092	1.82%
261630033	05-Jul-07	0.0099178	0.000208	2.10%
261630033	11-Jul-07	0.0089198	0.00001785	0.20%
261630033	17-Jul-07	0.0053044	0.0000369	0.70%
261630033	23-Jul-07	0.0095315	0.0000733	0.77%
261630033	29-Jul-07	0.0028393	0.0000218	0.77%
261630033	04-Aug-07	0.0036665	0.0000122	0.33%
261630033	10-Aug-07	0.004089	0.0000192	0.47%
261630033	16-Aug-07	0.0047761	0.0000143	0.30%
261630033	03-Sep-07	0.0039197	0.00001	0.26%
261630033	09-Sep-07	0.003146	0.0000459	1.46%
261630033	15-Sep-07	0.0026065	0.0000111	0.43%
261630033	21-Sep-07	0.0091525	0.0000933	1.02%
261630033	27-Sep-07	0.0046416	0.0000289	0.62%
261630033	03-Oct-07	0.005981	0.0000272	0.45%
261630033	06-Oct-07	0.005018	0.0000261	0.52%
261630033	09-Oct-07	0.006514	0.0000452	0.69%
261630033	21-Oct-07	0.0064791	0.0000363	0.56%
261630033	27-Oct-07	0.0036747	0.0000286	0.78%
261630033	02-Nov-07	0.011796	0.000104	0.88%
261630033	08-Nov-07	0.0073626	0.00006955	0.94%
261630033	14-Nov-07	0.00491865	0.0000085	0.17%
261630033	26-Nov-07	0.0087624	0.0000771	0.88%
261630033	02-Dec-07	0.00757165	0.000114	1.51%
261630033	08-Dec-07	0.00290345	0.0000052	0.18%
261630033	14-Dec-07	0.0058518	0.0000094	0.16%
261630033	20-Dec-07	0.0035523	0.0000265	0.75%
261630033	26-Dec-07	0.0074522	0.0000269	0.36%
261630033	01-Jan-08	0.0021103	0.0000207	0.98%
261630033	07-Jan-08	0.00776675	0.0000895	1.15%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
261630033	13-Jan-08	0.00263115	0.0000218	0.83%
261630033	19-Jan-08	0.0069021	0.0000154	0.22%
261630033	25-Jan-08	0.00689135	0.00002345	0.34%
261630033	31-Jan-08	0.0056087	0.0000118	0.21%
261630033	06-Feb-08	0.0020398	0.0000577	2.83%
261630033	18-Feb-08	0.0071664	0.0000163	0.23%
261630033	24-Feb-08	0.0068614	0.0000405	0.59%
261630033	07-Mar-08	0.00454905	0.0000129	0.28%
261630033	13-Mar-08	0.00700795	0.0000171	0.24%
261630033	19-Mar-08	0.0036055	0.0000262	0.73%
261630033	25-Mar-08	0.01249485	0.0000499	0.40%
261630033	06-Apr-08	0.00341975	0.0000068	0.20%
261630033	12-Apr-08	0.0036005	0.0000106	0.29%
261630033	18-Apr-08	0.01058655	0.000116	1.10%
261630033	24-Apr-08	0.0134431	0.0000972	0.72%
261630033	30-Apr-08	0.0082923	0.0000658	0.79%
261630033	06-May-08	0.00720835	0.0000643	0.89%
261630033	12-May-08	0.0039805	0.0000364	0.91%
261630033	18-May-08	0.0041587	0.0000177	0.43%
261630033	24-May-08	0.00435555	0.0000306	0.70%
261630033	30-May-08	0.01394445	0.000145	1.04%
261630033	05-Jun-08	0.01139835	0.000046	0.40%
261630033	11-Jun-08	0.00839205	0.0000337	0.40%
261630033	17-Jun-08	0.0060467	0.0000164	0.27%
261630033	23-Jun-08	0.007937	0.000105	1.32%
261630033	05-Jul-08	0.00944635	0.000392	4.15%
261630033	11-Jul-08	0.00784415	0.0000328	0.42%
261630033	17-Jul-08	0.01200845	0.0000669	0.56%
261630033	29-Jul-08	0.00836735	0.0000683	0.82%
261630033	10-Aug-08	0.00291315	0.0000049	0.17%
261630033	22-Aug-08	0.0073331	0.000075	1.02%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
261630033	28-Aug-08	0.00407345	0.0000306	0.75%
261630033	03-Sep-08	0.01224205	0.0000838	0.68%
261630033	09-Sep-08	0.0042144	0.000039	0.93%
261630033	15-Sep-08	0.00446255	0.0000181	0.41%
261630033	21-Sep-08	0.0030254	0.0000065	0.21%
261630033	27-Sep-08	0.00515685	0.0000353	0.68%
261630033	03-Oct-08	0.0052431	0.0000406	0.77%
261630033	09-Oct-08	0.0074004	0.0000188	0.25%
261630033	15-Oct-08	0.0133567	0.000132	0.99%
261630033	21-Oct-08	0.00618275	0.0000139	0.22%
261630033	27-Oct-08	0.00618555	0.0000099	0.16%
261630033	02-Nov-08	0.0022856	0.0000207	0.91%
261630033	08-Nov-08	0.0065656	0.0000225	0.34%
261630033	14-Nov-08	0.0060115	0.000111	1.85%
261630033	26-Nov-08	0.0037806	0.0000033	0.09%
261630033	02-Dec-08	0.0063334	0.0000084	0.13%
261630033	08-Dec-08	0.00391805	0.0000312	0.80%
261630033	26-Dec-08	0.00228905	0.0000216	0.94%
261630033	01-Jan-09	0.00302505	0.000372	12.30%
261630033	07-Jan-09	0.00194745	0.0000103	0.53%
261630033	13-Jan-09	0.0025347	0.0000142	0.56%
261630033	25-Jan-09	0.0037742	0.00001915	0.51%
261630033	31-Jan-09	0.00513625	0.0000219	0.43%
261630033	06-Feb-09	0.006904	0.0000278	0.40%
261630033	24-Feb-09	0.0040371	0.0000126	0.31%
261630033	08-Mar-09	0.0023407	0.0000288	1.23%
261630033	14-Mar-09	0.0053647	0.0000262	0.49%
261630033	20-Mar-09	0.0086397	0.0000885	1.02%
261630033	26-Mar-09	0.00506635	0.0000183	0.36%
261630033	01-Apr-09	0.0089369	0.0000405	0.45%
261630033	19-Apr-09	0.0020641	0.0000088	0.43%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
261630033	25-Apr-09	0.0110231	0.0000758	0.69%
261630033	07-May-09	0.0062342	0.000036	0.58%
261630033	13-May-09	0.0046832	0.000317	6.77%
261630033	19-May-09	0.0071426	0.0000865	1.21%
261630033	31-May-09	0.002204	0.00001925	0.87%
261630033	24-Jun-09	0.00577255	0.0000577	1.00%
261630033	30-Jun-09	0.0020965	0.0000181	0.86%
261630033	06-Jul-09	0.00457115	0.0000211	0.46%
261630033	12-Jul-09	0.0018438	0.0000082	0.44%
261630033	24-Jul-09	0.004156	0.0000392	0.94%
261630033	30-Jul-09	0.00377575	0.000048	1.27%
261630033	05-Aug-09	0.0058714	0.0000224	0.38%
261630033	11-Aug-09	0.0058233	0.0000173	0.30%
261630033	17-Aug-09	0.01014465	0.0000402	0.40%
261630033	04-Sep-09	0.00984245	0.000026	0.26%
261630033	10-Sep-09	0.0039165	0.0000054	0.14%
261630033	16-Sep-09	0.00372635	0.000024	0.64%
261630033	22-Sep-09	0.00403935	0.0000596	1.48%
261630033	28-Sep-09	0.0037259	0.0000374	1.00%
261630033	22-Oct-09	0.0089921	0.0000705	0.78%
261630033	28-Oct-09	0.0064514	0.0000317	0.49%
261630033	03-Nov-09	0.00428915	0.0000113	0.26%
261630033	09-Nov-09	0.00741965	0.0000516	0.70%
261630033	15-Nov-09	0.0033398	0.0000141	0.42%
261630033	21-Nov-09	0.00323275	0.000126	3.90%
261630033	27-Nov-09	0.00284575	0.0000145	0.51%
261630033	03-Dec-09	0.0023493	0.0000189	0.80%
261630033	09-Dec-09	0.0055543	0.0000711	1.28%
261630033	21-Dec-09	0.0028519	0.0000182	0.64%
261630033	14-Jan-10	0.0134654	0.000081	0.60%
261630033	20-Jan-10	0.0030915	0.0000177	0.57%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
261630033	26-Jan-10	0.0066768	0.00002165	0.32%
261630033	01-Feb-10	0.0064505	0.00005	0.78%
261630033	19-Feb-10	0.0037339	0.0000145	0.39%
261630033	25-Feb-10	0.0025884	0.0000187	0.72%
261630033	03-Mar-10	0.003534	0.0000192	0.54%
261630033	09-Mar-10	0.0115973	0.000138	1.19%
261630033	15-Mar-10	0.00270175	0.0000262	0.97%
261630033	21-Mar-10	0.0021123	0.0000091	0.43%
261630033	27-Mar-10	0.0024709	0.000015	0.61%
261630033	02-Apr-10	0.0092247	0.000125	1.36%
261630033	08-Apr-10	0.00361325	0.0000307	0.85%
261630033	14-Apr-10	0.00529515	0.0000546	1.03%
261630033	20-Apr-10	0.009515	0.000112	1.18%
261630033	26-Apr-10	0.00283225	0.0000313	1.11%
261630033	02-May-10	0.00503465	0.0000799	1.59%
261630033	14-May-10	0.0042193	0.0000438	1.04%
261630033	20-May-10	0.00542165	0.0000712	1.31%
261630033	26-May-10	0.0084684	0.0000535	0.63%
261630033	07-Jun-10	0.0041909	0.0000261	0.62%
261630033	13-Jun-10	0.00230345	0.0000477	2.07%
261630033	19-Jun-10	0.0042031	0.0000349	0.83%
261630033	25-Jun-10	0.00648495	0.0000512	0.79%
261630033	01-Jul-10	0.00425045	0.0000425	1.00%
261630033	07-Jul-10	0.00618485	0.0000499	0.81%
261630033	13-Jul-10	0.00563975	0.0000647	1.15%
261630033	19-Jul-10	0.0033244	0.000032	0.96%
261630033	25-Jul-10	0.0021857	0.00003195	1.46%
261630033	31-Jul-10	0.0045473	0.0000423	0.93%
261630033	06-Aug-10	0.0074914	0.0000206	0.27%
261630033	12-Aug-10	0.00444175	0.0000367	0.83%
261630033	18-Aug-10	0.0044524	0.0000289	0.65%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
261630033	24-Aug-10	0.00345765	0.0000323	0.93%
261630033	05-Sep-10	0.0045117	0.0000312	0.69%
261630033	11-Sep-10	0.003195	0.0000537	1.68%
261630033	17-Sep-10	0.00462125	0.000049	1.06%
261630033	29-Sep-10	0.0086121	0.0000878	1.02%
261630033	05-Oct-10	0.00424295	0.0000245	0.58%
261630033	11-Oct-10	0.0082004	0.0000446	0.54%
261630033	23-Oct-10	0.0102699	0.000154	1.50%
261630033	04-Nov-10	0.0058438	0.0000116	0.20%
261630033	10-Nov-10	0.0029787	0.0000343	1.15%
261630033	16-Nov-10	0.0072548	0.000099	1.36%
261630033	22-Nov-10	0.00786985	0.000124	1.58%
261630033	28-Nov-10	0.0046147	0.0000537	1.16%
261630033	04-Dec-10	0.002488	0.0000146	0.59%
261630033	10-Dec-10	0.00705425	0.000061	0.86%
261630033	16-Dec-10	0.0035897	0.0000415	1.16%
261630033	22-Dec-10	0.0018619	0.0000209	1.12%
261630033	28-Dec-10	0.00493515	0.0000324	0.66%
450250001	19-Jan-08	0.001	0.000018	1.80%
450250001	12-Feb-08	0.0005	0.0000039	0.78%
450250001	13-Mar-08	0.0015	0.0000064	0.43%
450250001	24-Apr-08	0.002	0.0000025	0.13%
450250001	30-Apr-08	0.0015	0.0000079	0.53%
450250001	18-May-08	0.0025	0.0000036	0.14%
450250001	30-May-08	0.003	0.0000051	0.17%
450250001	05-Jul-08	0.008	0.0000133	0.17%
450250001	23-Jul-08	0.0055	0.0000156	0.28%
450250001	29-Jul-08	0.003	0.000003	0.10%
450250001	10-Aug-08	0.001	0.0000046	0.46%
450250001	09-Sep-08	0.003	0.0000127	0.42%
450250001	15-Sep-08	0.002	0.0000098	0.49%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
510870014	03-Oct-08	0.00221	0.0000046	0.21%
510870014	14-Nov-08	0.00233	0.000008	0.34%
510870014	13-Jan-09	0.00312	0.000013	0.42%
510870014	20-Mar-09	0.00176	0.0000192	1.09%
510870014	06-Jul-09	0.00171	0.0000134	0.78%
510870014	12-Jul-09	0.0015	0.0000144	0.96%
510870014	24-Jul-09	0.00169	0.00001565	0.93%
510870014	16-Sep-09	0.00195	0.00000425	0.22%
510870014	21-Nov-09	0.00209	0.0000111	0.53%
510870014	03-Dec-09	0.00196	0.000008	0.41%
510870014	09-Dec-09	0.00226	0.000015	0.66%
510870014	20-Jan-10	0.00225	0.000026	1.16%
510870014	19-Feb-10	0.00276	0.000017	0.62%
510870014	09-Mar-10	0.00314	0.0000183	0.58%
510870014	21-Mar-10	0.00216	0.00000345	0.16%
510870014	02-Apr-10	0.00223	0.0000136	0.61%
510870014	08-Apr-10	0.00256	0.000019	0.74%
510870014	20-Apr-10	0.00242	0.0000197	0.81%
510870014	26-Apr-10	0.00198	0.0000229	1.16%
510870014	20-May-10	0.00205	0.0000236	1.15%
510870014	26-May-10	0.00208	0.0000267	1.28%
510870014	01-Jun-10	0.00175	0.0000209	1.19%
510870014	13-Jun-10	0.0017	0.0000363	2.14%
510870014	25-Jun-10	0.00156	0.00002	1.28%
510870014	07-Jul-10	0.0019	0.0000201	1.06%
510870014	13-Jul-10	0.00214	0.0000146	0.68%
510870014	19-Jul-10	0.00177	0.0000111	0.63%
510870014	25-Jul-10	0.00195	0.00001165	0.60%
510870014	31-Jul-10	0.00168	0.0000171	1.02%
510870014	06-Aug-10	0.00205	0.0000282	1.38%
510870014	12-Aug-10	0.00154	0.0000399	2.59%

AQS Site Code	Sample Date	Chromium TSP (µg/m^3)	Chromium VI (µg/m^3)	Chromium VI:Chromium TSP (percentage)
510870014	18-Aug-10	0.00142	0.0000266	1.87%
510870014	30-Aug-10	0.00178	0.0000178	1.00%
510870014	11-Sep-10	0.001626	0.0000114	0.70%
510870014	17-Sep-10	0.001835	0.000021	1.14%
510870014	23-Sep-10	0.002045	0.0000183	0.89%
510870014	11-Oct-10	0.002208	0.0000147	0.67%
510870014	10-Nov-10	0.001874	0.0000154	0.82%
510870014	16-Nov-10	0.001565	0.0000117	0.75%
510870014	22-Nov-10	0.002089	0.0000099	0.47%
510870014	04-Dec-10	0.001555	0.0000136	0.87%
510870014	10-Dec-10	0.00269	0.0000146	0.54%
510870014	16-Dec-10	0.001957	0.0000171	0.87%
510870014	22-Dec-10	0.001965	0.000017	0.87%
510870014	28-Dec-10	0.001719	0.0000153	0.89%
550270007	13-Oct-05	0.00155	0.00003645	2.35%
550270007	12-Nov-05	0.00136	0.0000117	0.86%

Appendix F

Model Evaluation Summaries

This appendix provides results of the model evaluation. For the pollutants modeled using the hybrid approach, we include here the information in Section 3.3 of the TSD and provide additional model performance statistics. The modeled and monitored values are provided in comma-separated value files in the folder "hybrid-polls-model_evaluation_paired" within the SupplementalData folder.

We describe here our efforts to evaluate the performance of the NATA models. Discussions of "HEM-3" in this document often are specifically related to the AERMOD dispersion model component of HEM-3, but we use "HEM-3" throughout for simplicity and consistency. In this section in particular, discussions of HEM-3 model values are specifically related to the air concentrations predicted by its AERMOD component.

F.1 Overview

Using the air toxics archive Phase IX for the year 2011, we conducted an operational model performance evaluation of the air toxics simulated for the 2011 NATA (more details found in Section F.2 below). The model evaluation included both the air toxics modeled with the hybrid approach ("hybrid air toxics") and those modeled without the hybrid approach ("non-hybrid air toxics"). The hybrid evaluation looked at the air toxics for which there were valid ambient data (i.e., completeness criteria protocol) to compare against the CMAQ, HEM-3, and hybrid model predictions. Likewise, the air toxics non-hybrid evaluation used similar observational-completeness criteria constraints to compare against air toxics estimated by adding HEM-3 to observed ambient concentrations assumed to reflect background conditions.

Spatial-scale differences exist between CMAQ, HEM-3, and the hybrid model predictions. A CMAQ concentration represents a 12-km grid-cell volume-averaged value. The HEM-3 model concentration represents a specific point within the modeled domain. The hybrid model concentration combines the HEM-3 point-concentration gradients with the CMAQ 12-km grid-cell volume average. The ambient observed measurements were made at specific spatial locations (latitude/longitude). Several annual graphical presentations and statistics of model performance were calculated and prepared. Graphical presentations included box and whisker plots (which show the distribution and the bias of the predicted and observed data) and regional maps (which show the mean bias and error calculated at individual monitoring sites).

F.2 Observations

Observations were extracted from the air toxics archive, Phase IX for the year 2011. While most of the data in the archive are a snapshot of the Air Quality System (AQS) database (downloaded in July 2014), additional data (such as from special studies) were in the archive but not reported to AQS. In the air toxics archive, pollutant concentrations were converted to µg/m^3 in local conditions where temperature and pressure data were available (i.e., at the vast majority of sites). In addition, any negatives and data flagged as "non-detect" without a value were given a value of 0. Also, any data determined to have been substituted with half the method detection limit (i.e., MDL/2) was changed to 0.

For comparing annual averages of modeled and monitored data, data from the archive were aggregated to 2011 annual averages by site and parameter code. Data below MDLs were used as-is. Data were removed for which there were no MDLs. Naphthalene data from parameter code 45850 (canister method) were removed because that method may not be as reliable as the method used in the National Ambient Air Toxics Trends Program. Also, those sites were removed that reported naphthalene as code 17141 in which it was determined to use the canister test method.

Only site-parameter pairs in which measurements from at least three seasons were 75 percent complete (i.e., 75 percent of the scheduled days contained non-null values) were retained for developing annual averages. First, the sub-annual data were allocated to 24-hour averages. Seventy-five-percent completeness was required to create a daily average from sub-daily data, such that 75 percent of scheduled sub-annual data were available. For example, hourly data required 18 of 24 hours of data, three-hour data required six of eight three-hour periods of data, etc. For each quarter, the number of days to meet 75% completeness depended on the sampling frequency (note that more than one monitor at the site that measured on the same day was counted once). For example, one-in-six-day sampling required 12 days for the quarter.

For sites with multiple monitors (known as "POCS"), only the daily data with the same measurement duration (i.e., hourly, 3-hour, 24-hour) were averaged across the POCs. That is, daily data based on hourly measurements were not averaged with daily data based on 24-hour measurements. Where a site met the 75% completeness for multiple durations, the 24-hour duration data were chosen.

Annual averages were created by averaging all daily measurements with the same measurement duration for all sites that met the above completeness criteria. The only sites used were those for which 50 percent or more of the data were above the method detection limit (MDL). A spreadsheet file ("2011monitored_data_annualmeans_PhaseIXarchive.xlsx") of the ambient annual averages (in µg/m^3) is provided in the SupplementalData folder.

Uncertainties in the ambient data result from limited sites, data below MDL and measurement uncertainties.

F.3 Model Performance Statistics

The Atmospheric Model Evaluation Tool (AMET) was used to conduct the 2011 NATA air toxics evaluation (Appel et al. 2011[1]). There are various statistical metrics available and used by the science community for model performance evaluation. For a robust evaluation, the principal evaluation statistics used to evaluate model performance are based on the following metrics: two bias metrics (mean bias and normalized mean bias); and, three error metrics (mean error and normalized mean error, root mean square error, and correlation coefficient).

Common variables are:

M = predicted concentration
O = observed concentration
X = predicted or observed concentration
σ = standard deviation

Mean Bias (MB), Mean Error (ME), and Root Mean Square Error (RMSE) (all in µg/m^3):

$$MB = \frac{1}{n}\sum_{1}^{n}(M - O)$$

$$ME = \frac{1}{n}\sum_{1}^{n}|M - O|$$

$$RMSE = \sqrt{\frac{\sum_{1}^{n}(M - O)^2}{n}}$$

MB quantifies the tendency of the model to over- or under-estimate values while ME and RMSE measure the magnitude of the difference between modeled and observe values regardless of whether the modeled values are higher or lower than observations.

Normalized Mean Bias (NMB) and Normalized Mean Error (NME) (both unitless):

$$NMB = \frac{\sum_{1}^{n}(M - O)}{\sum_{1}^{n} O}$$

---

[1] Appel, K.W., Gilliam, R.C., Davis, N., Zubrow, A., and Howard, S.C. 2011. Overview of the Atmospheric Model Evaluation Tool (AMET) v1.1 for Evaluating Meteorological and Air Quality Models. Environ. Modell. Softw., 26(4): 434–443.

NMB is used as a normalization to facilitate a range of concentration magnitudes. This statistic averages the difference (model - observed) over the sum of observed values. NMB is a useful model performance indicator because it avoids over-inflating the observed range of values, especially at low concentrations.

$$NME = \frac{\sum_1^n |M - O|}{\sum_1^n O}$$

NME is similar to NMB, where the performance statistic is used as a normalization of the ME. NME indicates the absolute value of the difference (model - observed) over the sum of observed values.

Correlation Coefficient (r; unitless):

$$r = \frac{1}{(n-1)} \sum_1^n \left[\left(\frac{O - \bar{O}}{\sigma_o} \right) \times \left(\frac{M - \bar{M}}{\sigma_m} \right) \right]$$

The value of r provides an indication of the strength of linear relationship and is signed positive or negative based on the slope of the linear regression.

Fractional Bias (FB) and Fractional Error (FE) (both unitless):

$$FB = \frac{1}{n} \left(\frac{\sum_1^n (M - O)}{\sum_1^n \left(\frac{(M + O)}{2} \right)} \right)$$

FB is a useful model performance indicator because it has the advantage of equally weighting positive and negative bias estimates. The single largest disadvantage in this estimate of model performance is that the estimated concentration (i.e., M) is found in both the numerator and denominator.

$$FE = \frac{1}{n} \left(\frac{\sum_1^n |M - O|}{\sum_1^n \left(\frac{(M + O)}{2} \right)} \right)$$

FE is similar to FB except the absolute value of the difference is used to that the error is always positive.

Standard Deviation (σ, ppb):

$$\sigma = \sqrt{\frac{1}{n} \sum \left(X - \bar{\bar{X}} \right)^2}$$

σ is a measure of the amount of variation of the observed and predicted values.

Coefficient of Variation (CoV, unitless):

$$CoV = \frac{\sigma}{\bar{\bar{X}}}$$

CoV is the ratio of σ to the mean and shows the extent of variation in relation to the mean.

Index of Agreement (unitless):

$$Index\ of\ Agreement = 1 - \left[\frac{\sum_1^n (O - M)^2}{\sum_1^n \left(\left| M - \bar{\bar{O}} \right| + \left| O - \bar{O} \right| \right)^2} \right]$$

Index of Agreement provides a sense of the strength of the relationship between model estimates and observations that have been paired in time and space.

Systematic RMSE ($RMSE_s$) and Unsystematic RMSE ($RMSE_u$) (both ppb):

$$RMSE_s = \sqrt{\frac{1}{n}\sum_{1}^{n}(C^* - C_o)^2}$$

$$RMSE_u = \sqrt{\frac{1}{n}\sum_{1}^{n}\left(C^* - C_p\right)^2}$$

Where:

$$C^* = a + bC_o$$

$$a = least\ squares\ regression\ coefficient\ of\ C_p$$

$$b = least\ squares\ regression\ coefficient\ of\ C_o$$

$$C_p = predicted\ (modeled)\ concentration$$

$$C_o = observed\ concentration$$

$RMSE_s$ measures the difference between the regression line of the observed and predicted values, while $RMSE_u$ measures the random error about the regression line of the predicted values.

Skewness (unitless):

Measures the asymmetry of the probability distribution of a random value about its mean. For this assessment, skewness was calculated as simply median/mean.

F.4 Hybrid Evaluation

We conducted an annual operational model performance evaluation for hybrid air toxics, resulting in comparisons between CMAQ and HEM-3 predictions as well as an evaluation of the ability of the hybrid model to replicate the 2011 observed ambient concentrations. Inclusion of all three model results was intended to demonstrate the merged attributes of the hybrid model used for the 2011 NATA. Statistical assessments of modeled versus observed concentrations were paired in time and space and aggregated on an annual basis. Exhibit F-1 contains a list of air toxics evaluated in the hybrid model performance evaluation and the number of paired sites (based on completeness criteria of observations, Section F-2) used in the annual average. Exhibit F-2 is a map of the 2011 monitoring locations for hybrid air toxics. Acrolein and ethylene dibromide were excluded in the model evaluation given the data uncertainty and sampling. Annual averages of xylene species (m-, o-, and p-) were summed together to calculate a "Xylenes, total" air toxic group based on the individual risk for each species being the same.

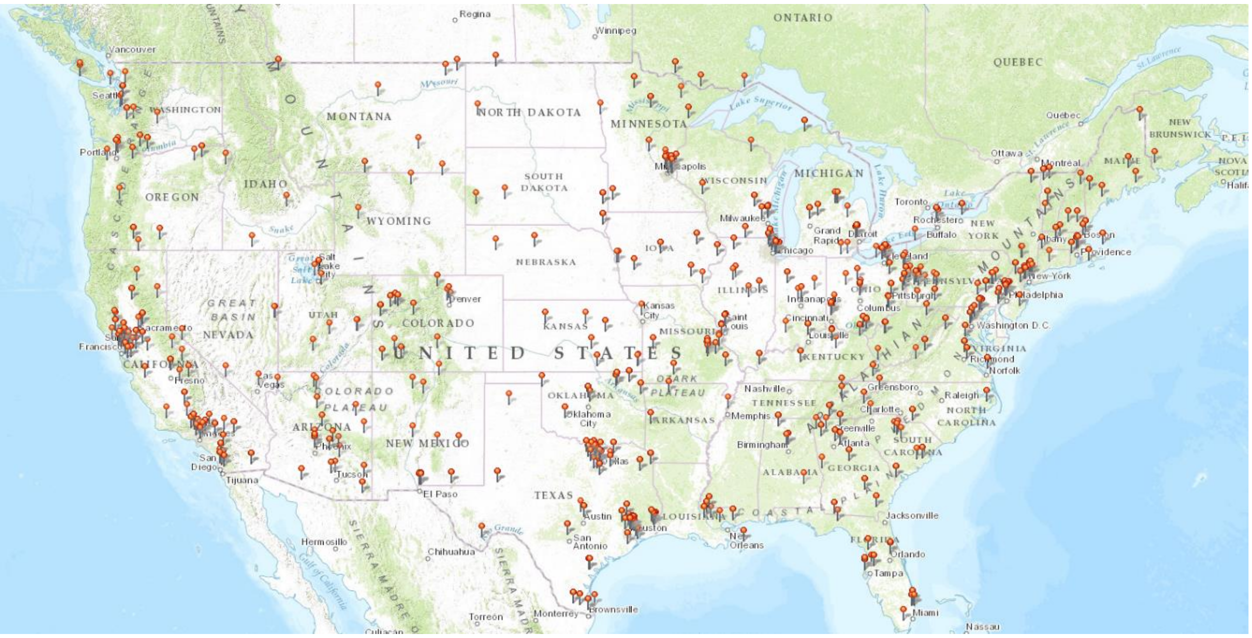
Exhibit F-1. Hybrid Air Toxics Evaluated

Hybrid Air Toxic	Number of Paired Sites
1,3-Butadiene	83
1,3-Dichloropropene	5
1,4-Dichlorobenzene	22
Acetaldehyde	110
Acrylonitrile	18
Formaldehyde	110
Lead PM_{10}	33
Manganese PM_{10}	40
Methylene chloride	123
Naphthalene	36

Hybrid Air Toxic	Number of Paired Sites
Arsenic PM_{10}	34
Benzene	214
Cadmium PM_{10}	27
Chlorine	123
Chloroform	92
Ethylene dichloride	40

Hybrid Air Toxic	Number of Paired Sites
Nickel PM_{10}	29
Propylene dichloride	5
Tetrachloroethylene	72
Toluene	211
Trichloroethylene	13
Xylenes	163

Exhibit F-2. 2011 Monitoring Locations for the Evaluation of Hybrid Air Toxics



The annual model performance results for seven of these key hybrid air toxics are presented below in Exhibit F-3.

Exhibit F-3. 2011 Annual Air Toxics Performance Statistics for the Hybrid, CMAQ, and HEM-3 Models

Hybrid Air Toxic	Model	MB (µg/m^3)	ME (µg/m^3)	NMB (%)	NME (%)
Acetaldehyde	Hybrid	0.5	0.7	30.9	43.9
	CMAQ	0.4	0.7	27.1	41.4
	HEM-3	-1.3	1.3	-80.4	80.6
Formaldehyde	Hybrid	-0.8	1.0	-30.8	37.3
	CMAQ	-0.9	1.0	-34.2	38.8
	HEM-3	-2.2	2.2	-78.8	79.3
Benzene	Hybrid	0.0	0.5	1.6	60.7
	CMAQ	-0.2	0.5	-22.7	57.8
	HEM-3	-0.3	0.5	-33.1	60.2

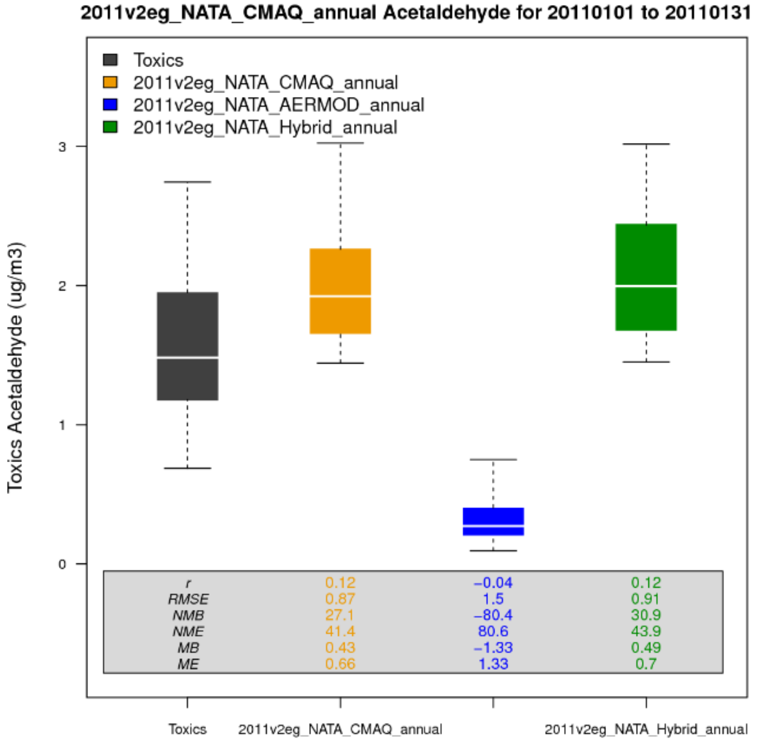
Hybrid Air Toxic	Model	MB (µg/m^3)	ME (µg/m^3)	NMB (%)	NME (%)
1,3-butadiene	Hybrid	0.0	0.1	1.6	78.5
	CMAQ	0.0	0.1	-21.2	77.2
	HEM-3	0.0	0.1	5.4	77.3
Toluene	Hybrid	1.4	2.0	91.9	128.0
	CMAQ	0.8	1.6	52.8	103.0
	HEM-3	1.2	1.8	74.9	116.0
Lead PM$_{10}$	Hybrid	0.0	0.0	48.1	208.0
	CMAQ	0.0	0.0	-10.4	152.0
	HEM-3	0.0	0.0	-11.2	168.0
Arsenic PM$_{10}$	Hybrid	0.0	0.0	-74.9	79.0
	CMAQ	0.0	0.0	-80.2	81.8
	HEM-3	0.0	0.0	-81.8	85.6

Boxplots showing model distribution (units of µg/m^3) and bias differences (units of µg/m^3) as compared to ambient observations are presented below. The boxplots use boxed interquartile ranges of 25th-to-75th percentile, along with whiskers from the 5th to 95th percentiles, and they also contain summary statistics of r, RMSE, NMB, NME, MB, and ME. Likewise, regional spatial maps which show the mean bias and error calculated at individual

As evidenced by Exhibits F-4 through F-24, CMAQ and hybrid model predictions of annual formaldehyde, acetaldehyde, and benzene (three key air toxics in NATA) showed relatively small-to-moderate bias and error percentages when compared to observations. HEM-3 showed larger biases and errors, with underestimates for secondarily formed air toxics (e.g., -80.4 percent for acetaldehyde and -78.8 percent for formaldehyde), as expected given HEM-3's exclusion of atmospheric chemistry. Differences in bias and error statistics between the hybrid and CMAQ models were negligible for formaldehyde and acetaldehyde. Technical issues in the air toxics data consisted of (1) uncertainties in monitoring methods, (2) limited measurements in time/space to characterize ambient concentrations ("local in nature"), (3) commensurability issues between measurements and model predictions, (4) emissions- and science-uncertainty issues potentially affecting model performance, and (5) limited data for estimating intercontinental transport that effects the estimation of boundary conditions (i.e., boundary estimates for some species were much higher than predicted values inside the domain).

Exhibits F-25 through F-53 contain the box plots and regional spatial maps for the remaining evaluated hybrid air toxics.

Exhibit F-4. Acetaldehyde: 2011 Boxplots of Observed and Modeled Concentrations (top) and Modeled-Observed Bias Difference in Concentrations (bottom) for the Hybrid, CMAQ, and HEM-3 Models


2011v2eg_NATA_CMAQ_annual Acetaldehyde for 20110101 to 20110131

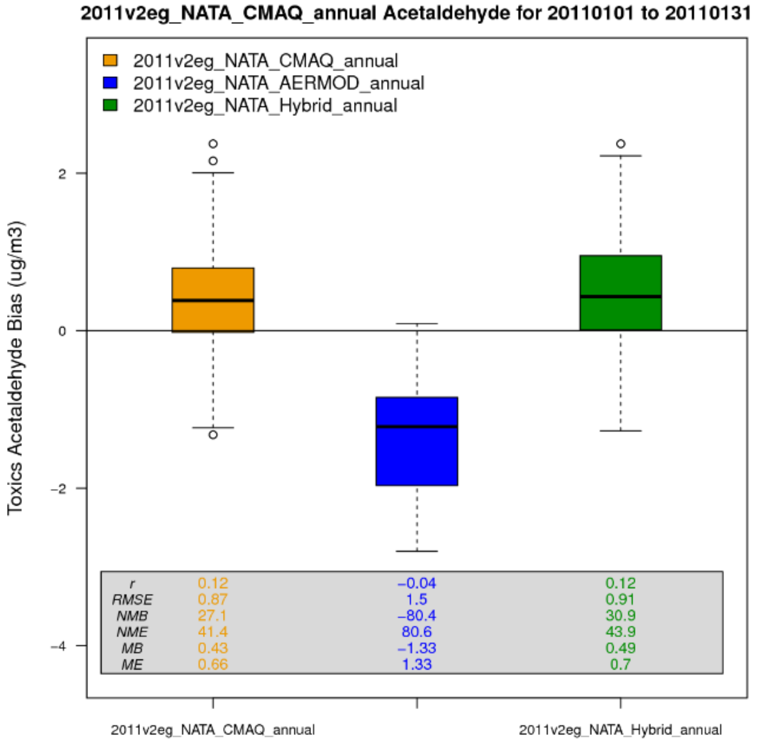

2011v2eg_NATA_CMAQ_annual Acetaldehyde for 20110101 to 20110131

Exhibit F-5. Acetaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the Hybrid Modeling Domain

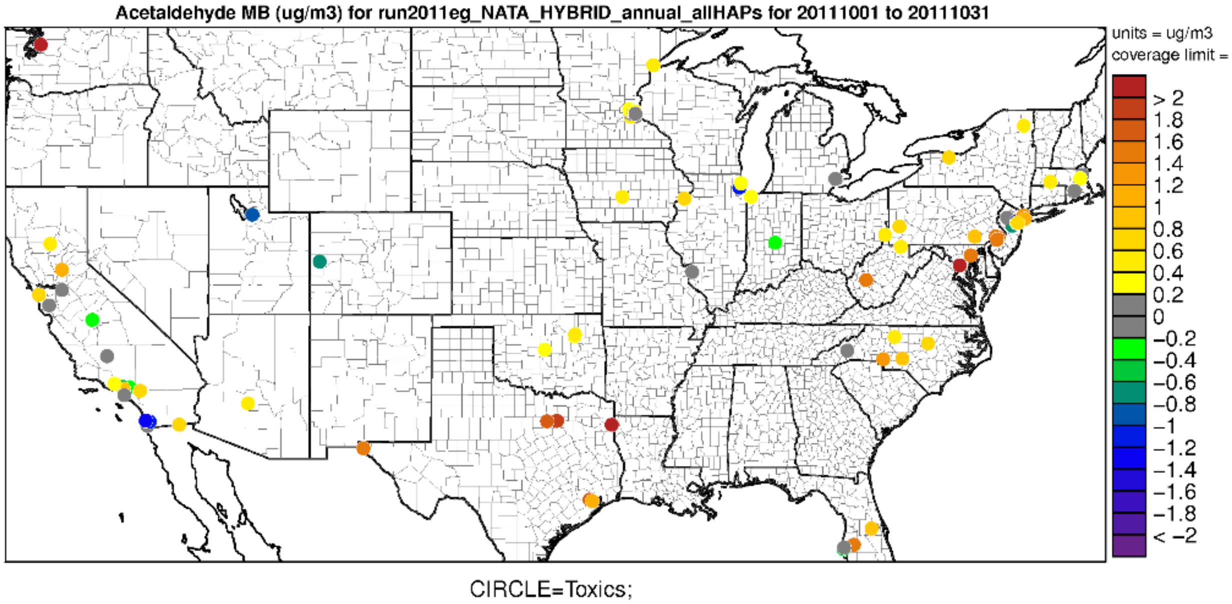


Exhibit F-6. Acetaldehyde: 2011 Mean Error (%) at Monitoring Sites in the Hybrid Modeling Domain

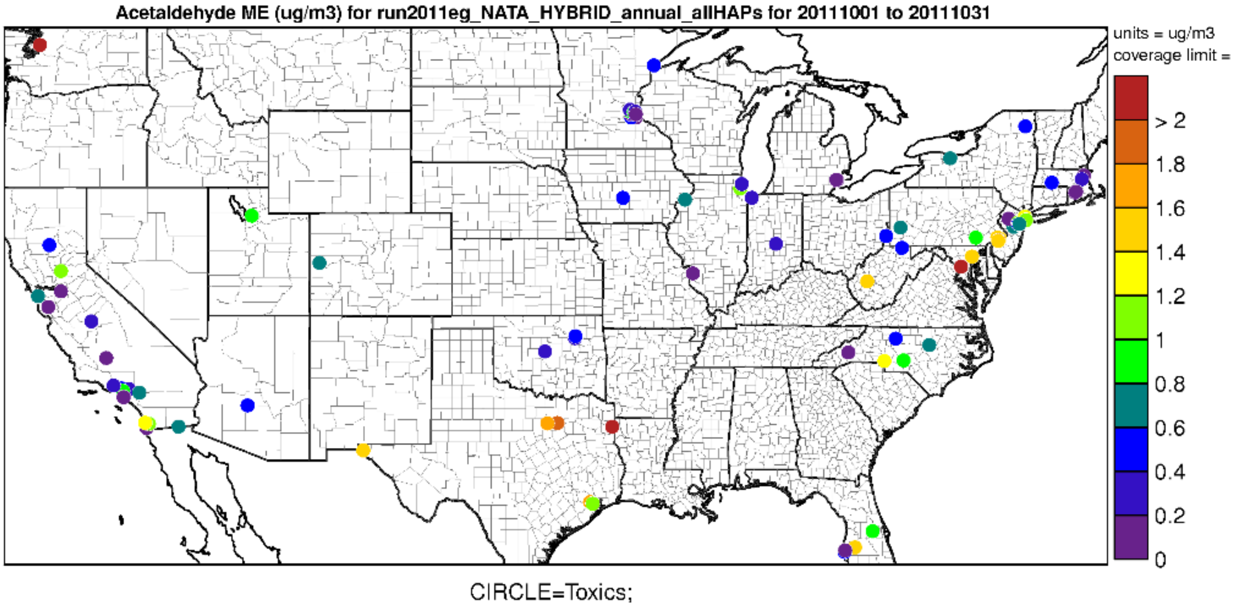


Exhibit F-7. Acetaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the CMAQ Modeling Domain

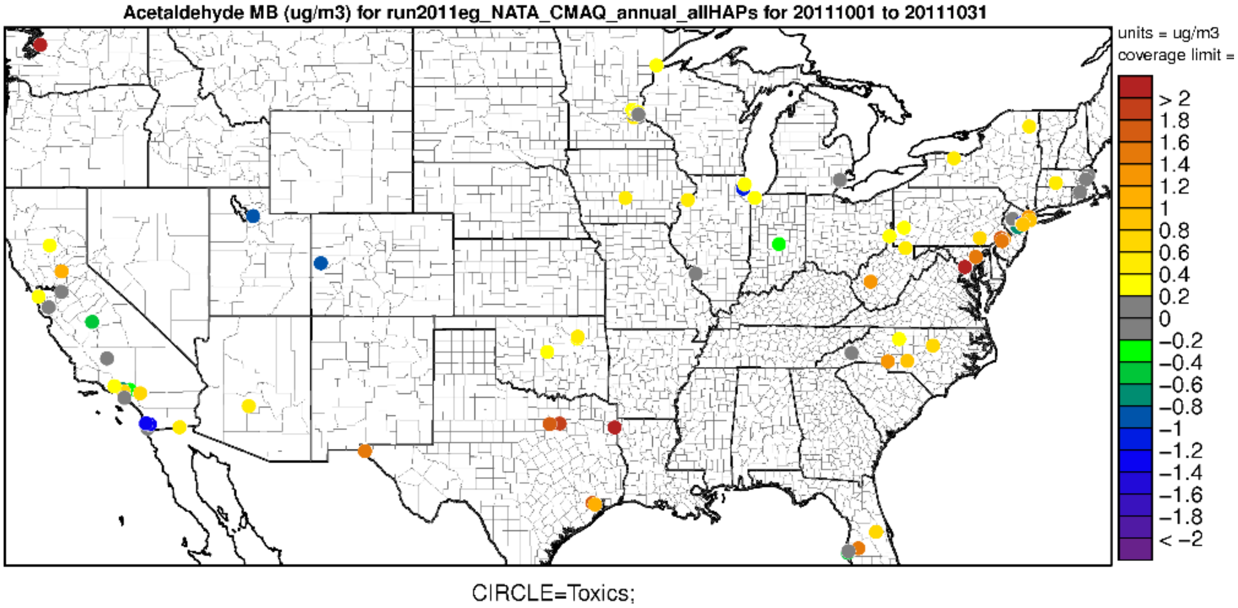


Exhibit F-8. Acetaldehyde: 2011 Mean Error (%) at Monitoring Sites in the CMAQ Modeling Domain

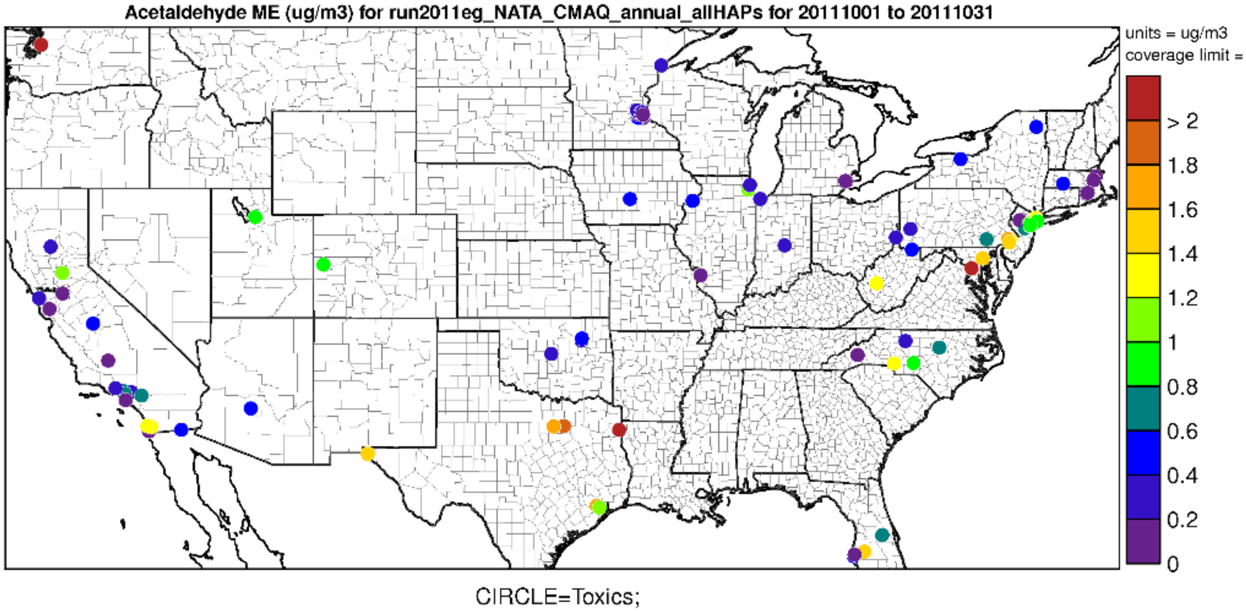


Exhibit F-9. Acetaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the HEM-3 Modeling Domain

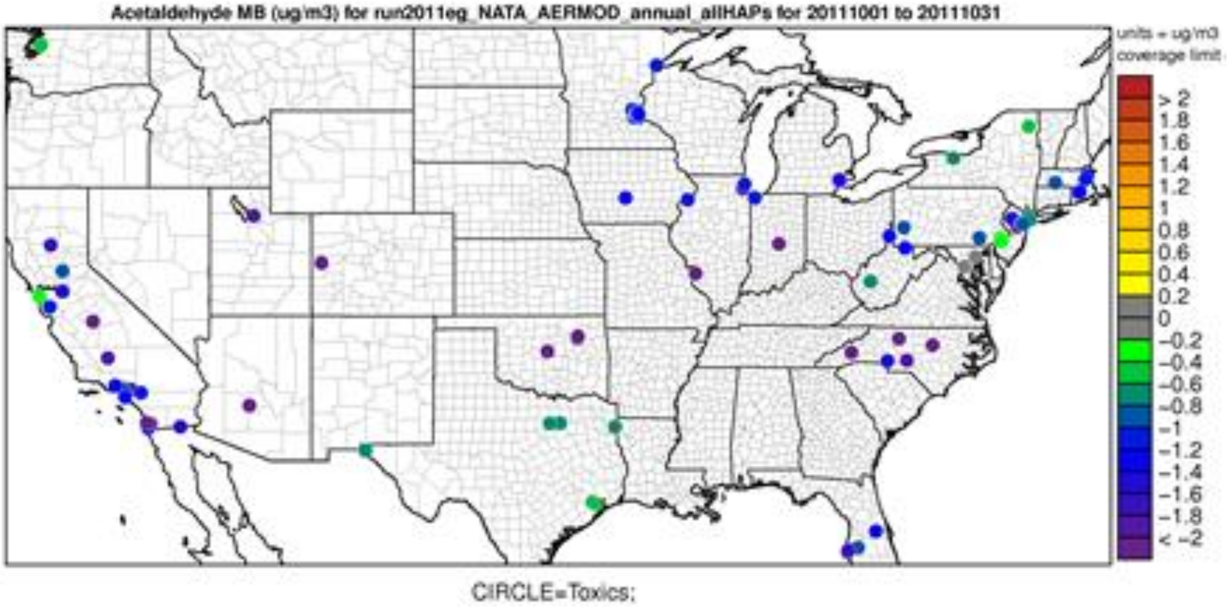


Exhibit F-10. Acetaldehyde: 2011 Mean Error (%) at Monitoring Sites in the HEM-3 Modeling Domain

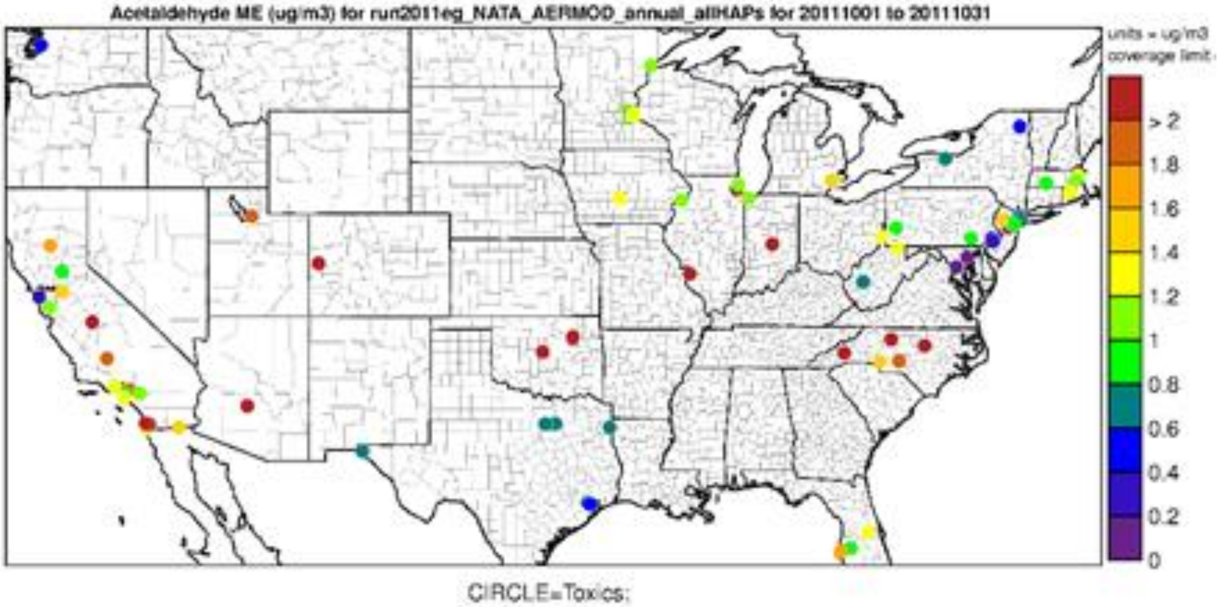
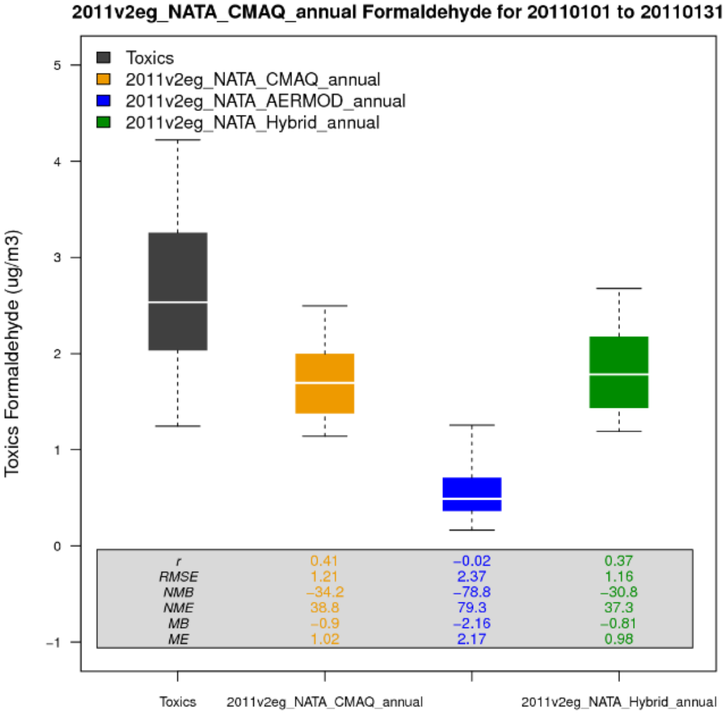


Exhibit F-11. Formaldehyde: 2011 Boxplots of Observed and Modeled Concentrations (top) and Modeled-Observed Bias Difference in Concentrations (bottom) for the Hybrid, CMAQ, and HEM-3 Models



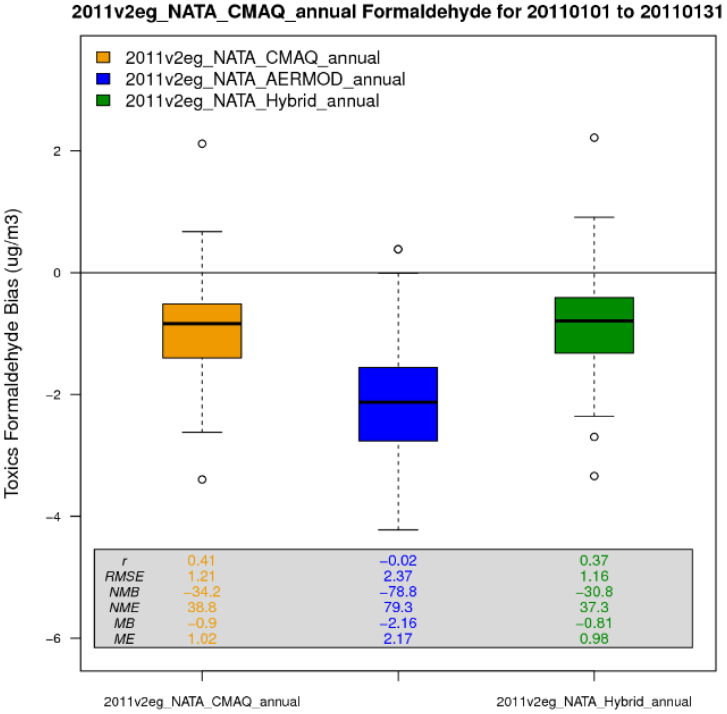


Exhibit F-12. Formaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the Hybrid Modeling Domain

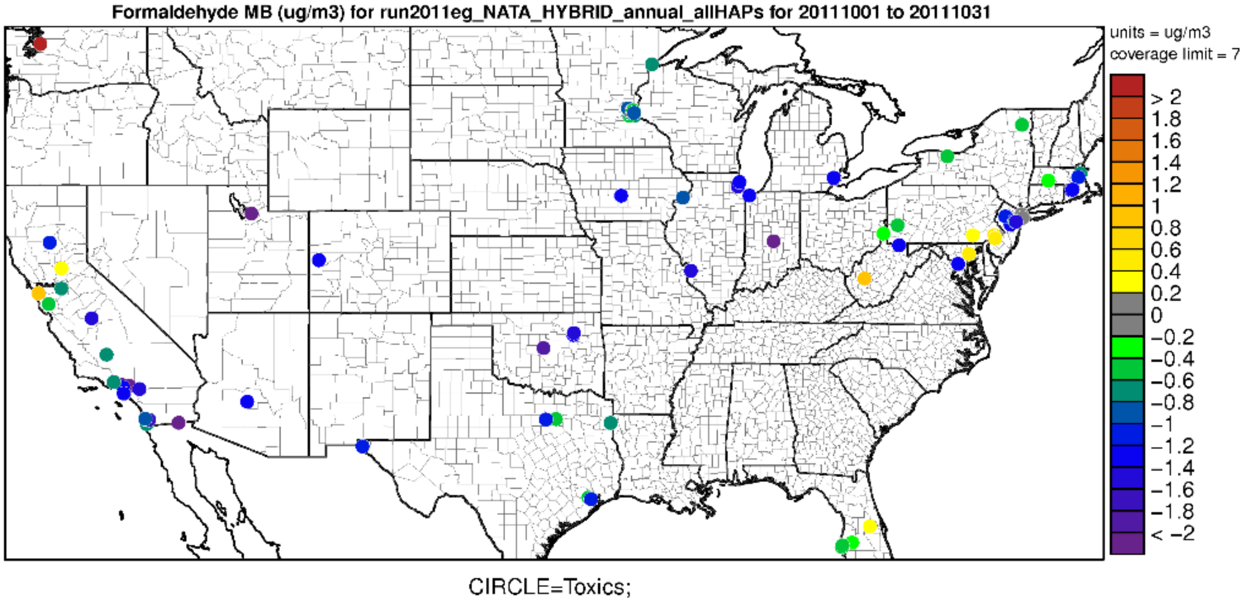


Exhibit F-13. Formaldehyde: 2011 Mean Error (%) at Monitoring Sites in the Hybrid Modeling Domain

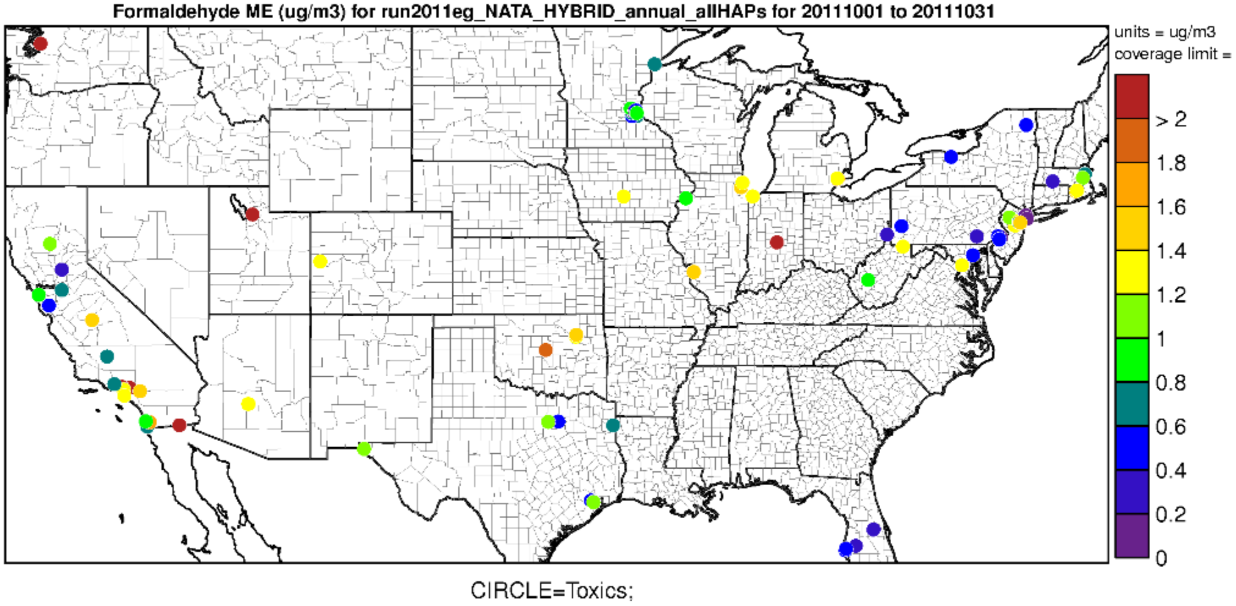


Exhibit F-14. Formaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the CMAQ Modeling Domain

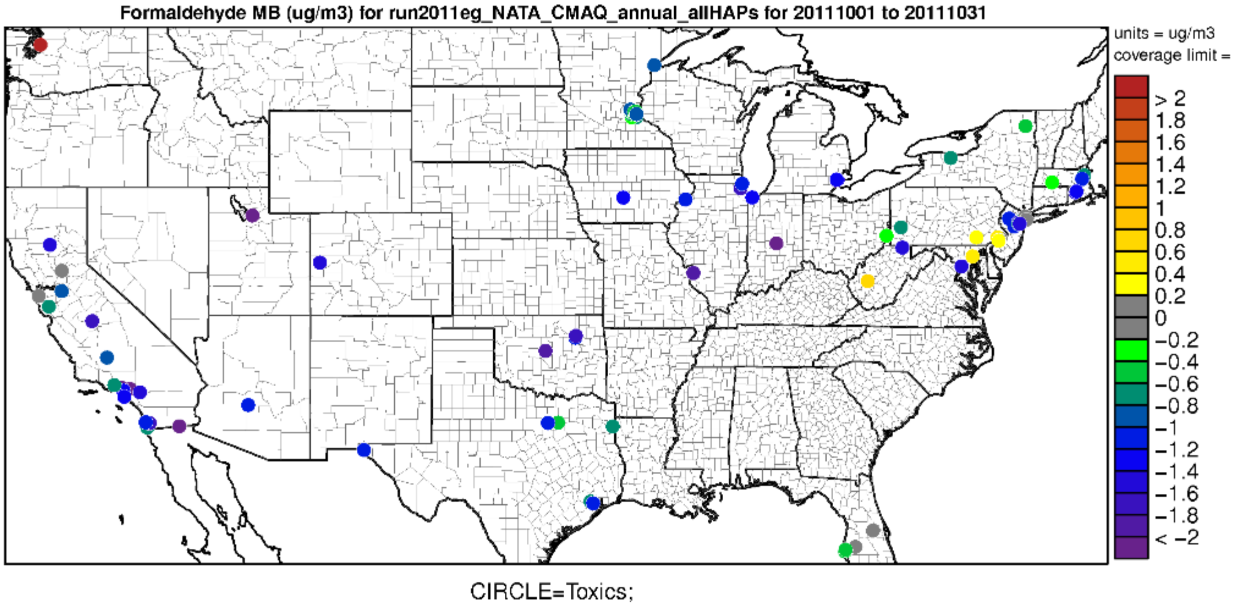


Exhibit F-15. Formaldehyde: 2011 Mean Error (%) at Monitoring Sites in the CMAQ Modeling Domain

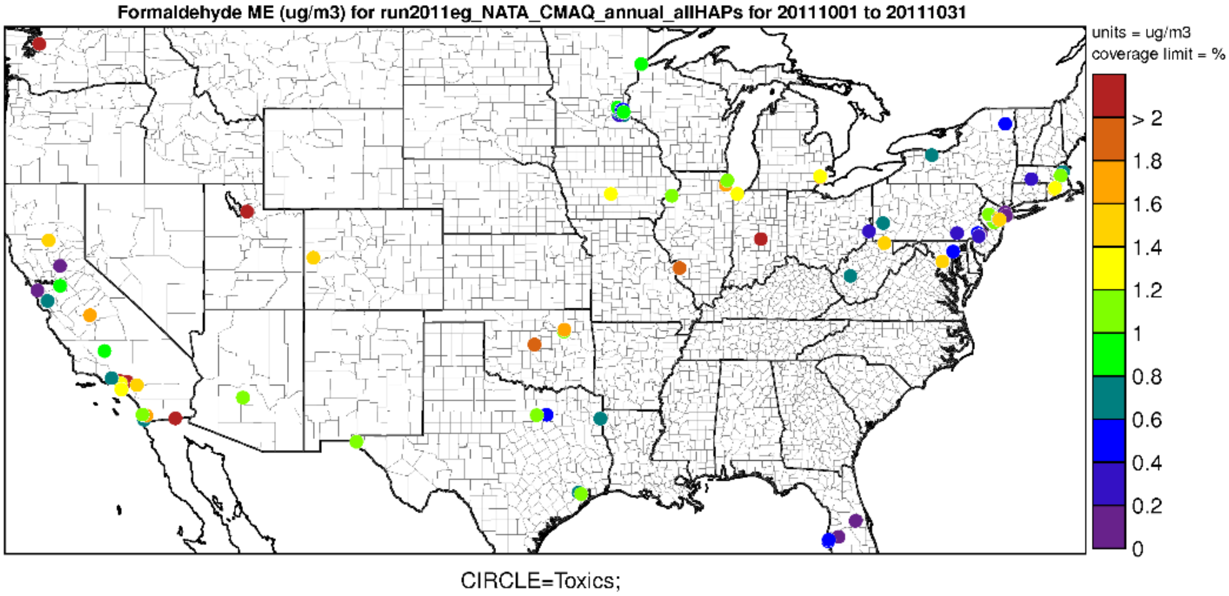


Exhibit F-16. Formaldehyde: 2011 Mean Bias (%) at Monitoring Sites in the HEM-3 Modeling Domain

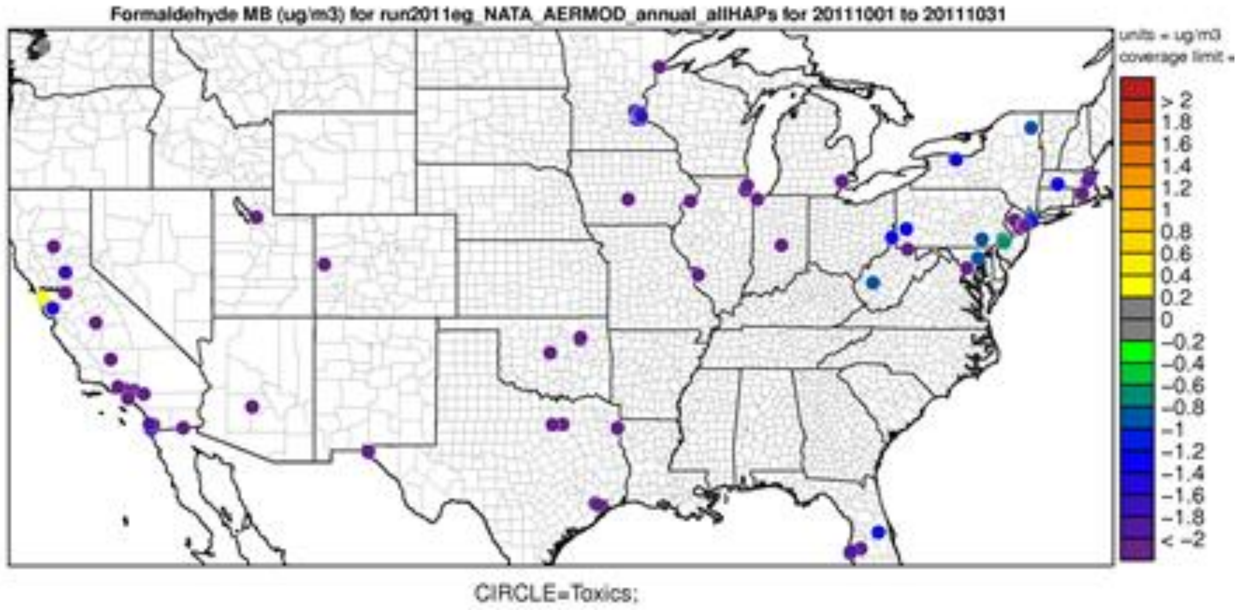


Exhibit F-17. Formaldehyde: 2011 Mean Error (%) at Monitoring Sites in the HEM-3 Modeling Domain

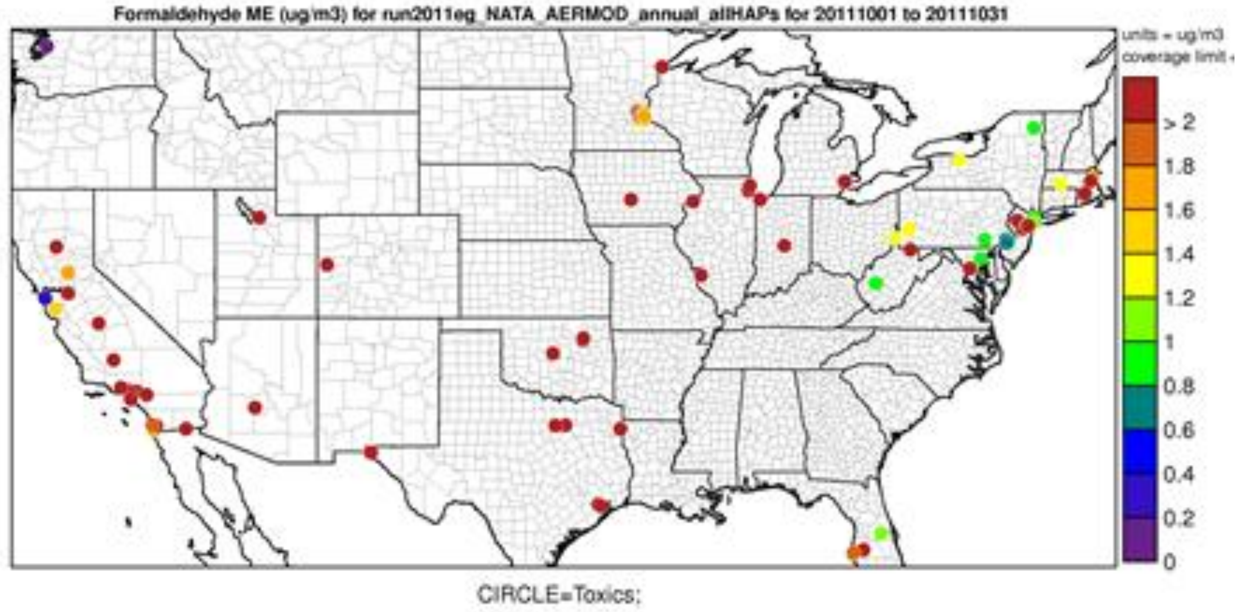
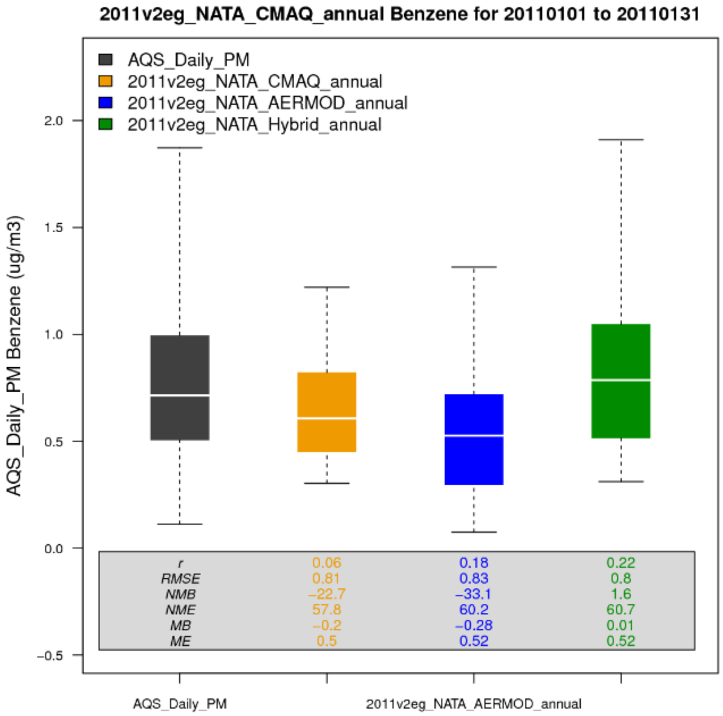


Exhibit F-18. Benzene: 2011 Boxplots of Observed and Modeled Concentrations (top) and Modeled-Observed Bias Difference in Concentrations (bottom) for the Hybrid, CMAQ, and HEM-3 Models



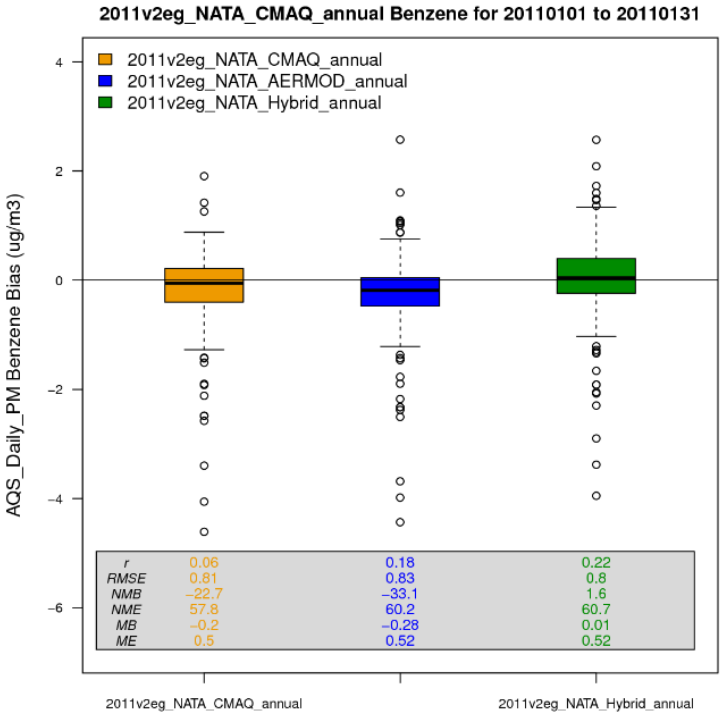


Exhibit F-19. Benzene: 2011 Mean Bias (%) at Monitoring Sites in the Hybrid Modeling Domain

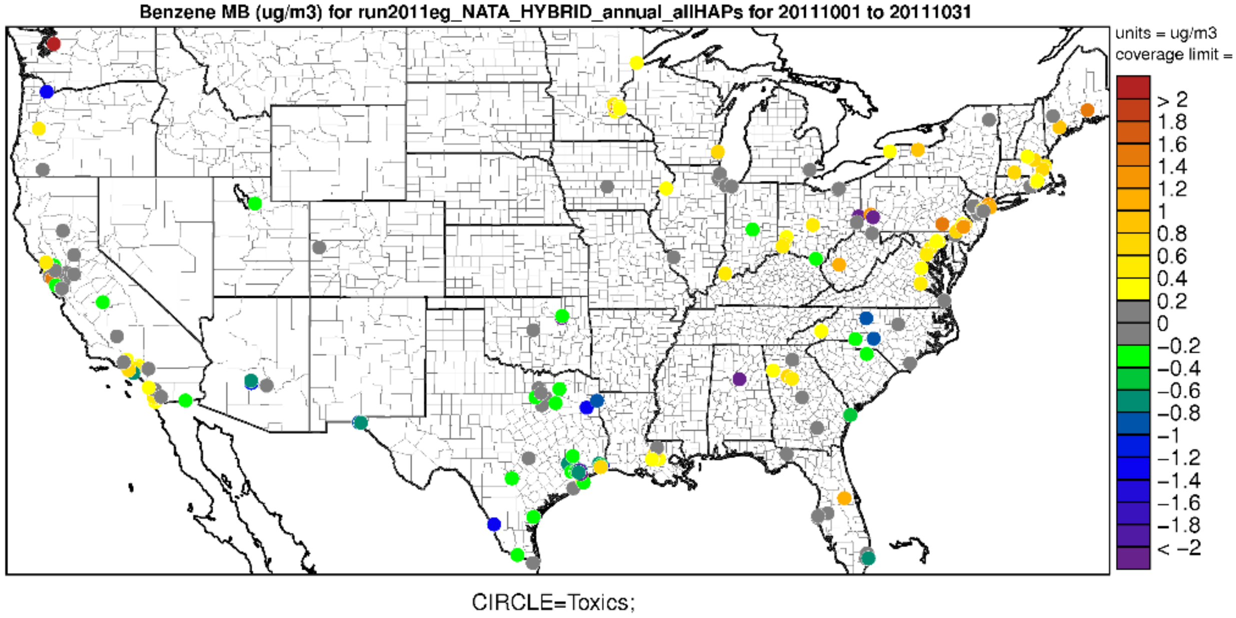


Exhibit F-20. Benzene: 2011 Mean Error (%) at Monitoring Sites in the Hybrid Modeling Domain

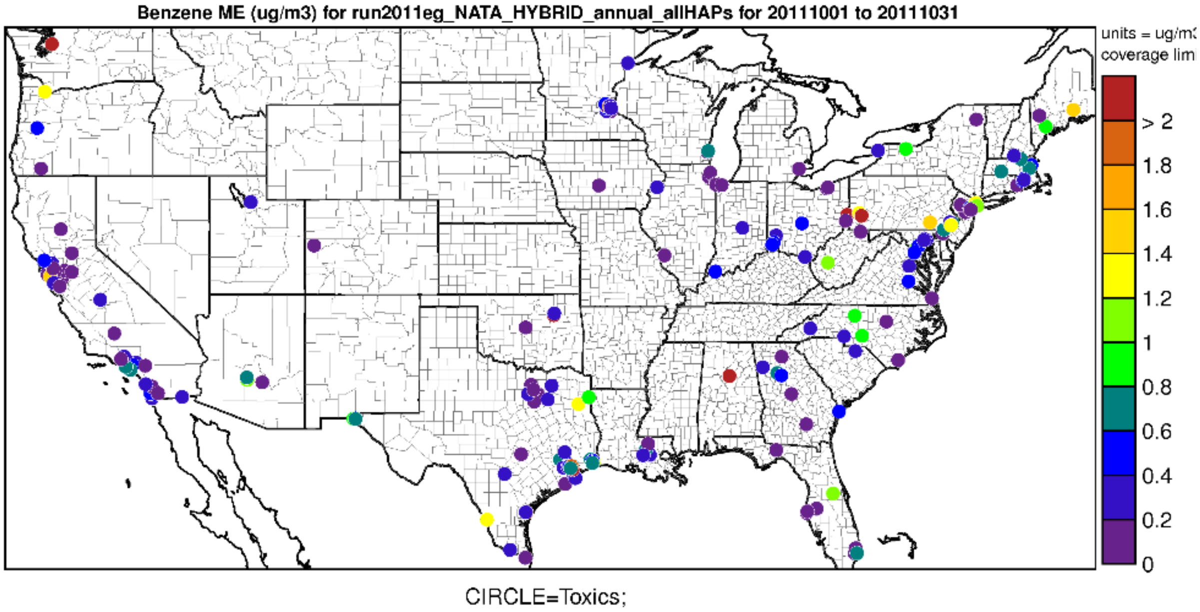


Exhibit F-21. Benzene: 2011 Mean Bias (%) at Monitoring Sites in the CMAQ Modeling Domain

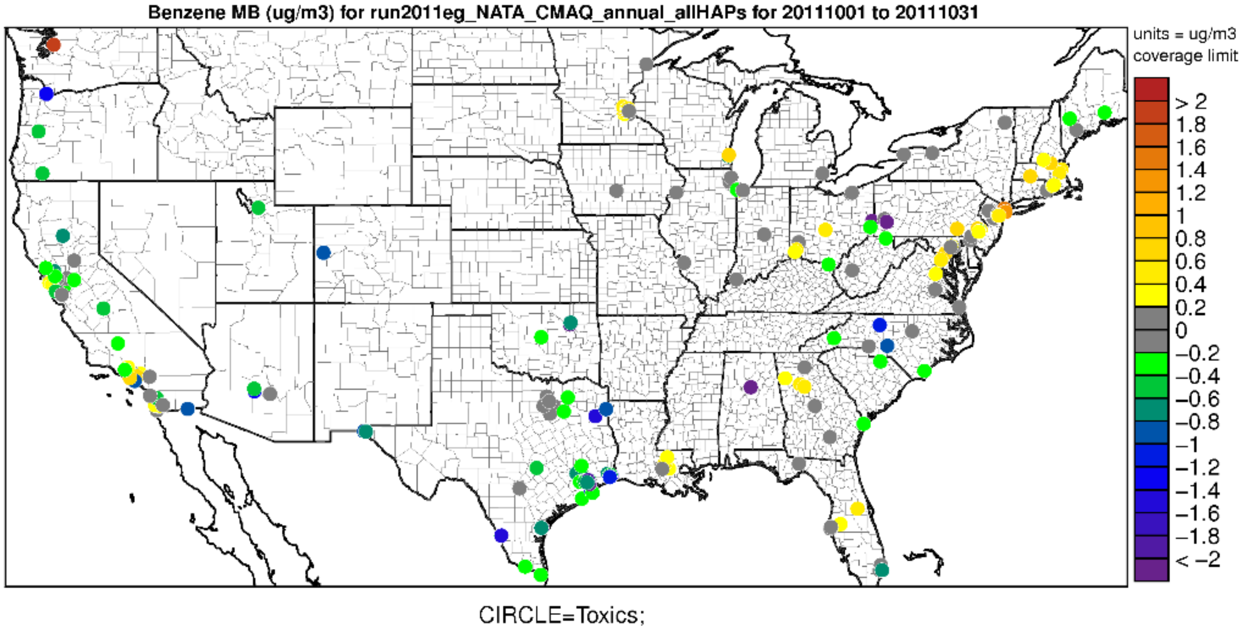


Exhibit F-22. Benzene: 2011 Mean Error (%) at Monitoring Sites in the CMAQ Modeling Domain

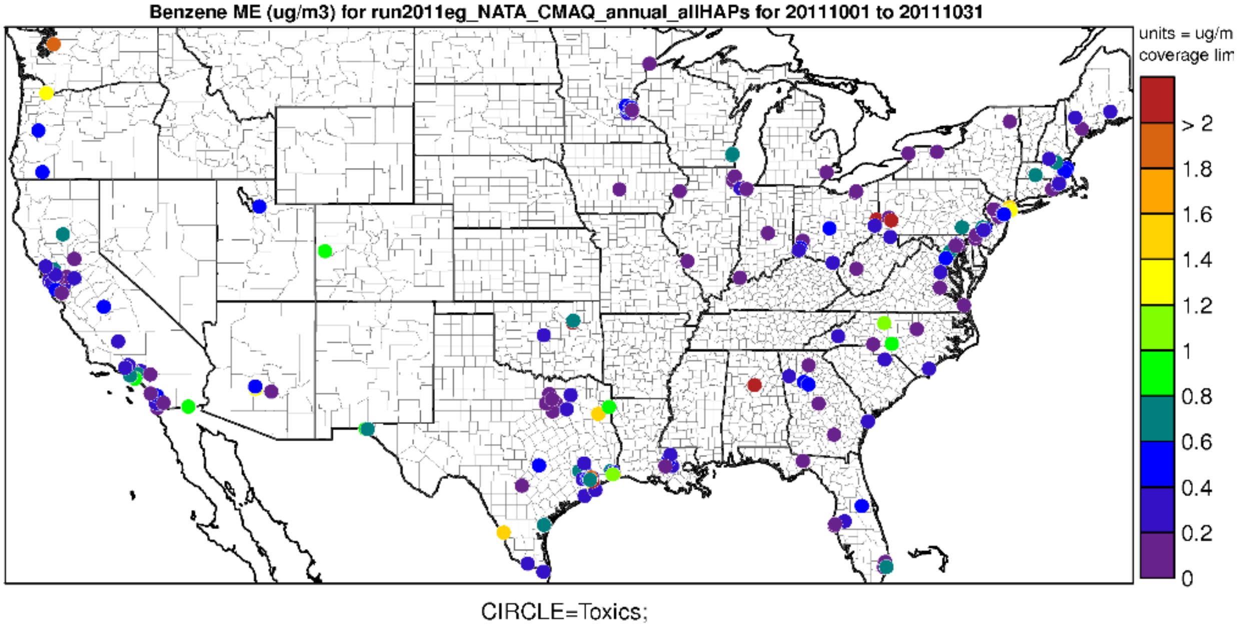


Exhibit F-23. Benzene: 2011 Mean Bias (%) at Monitoring Sites in the HEM-3 Modeling Domain

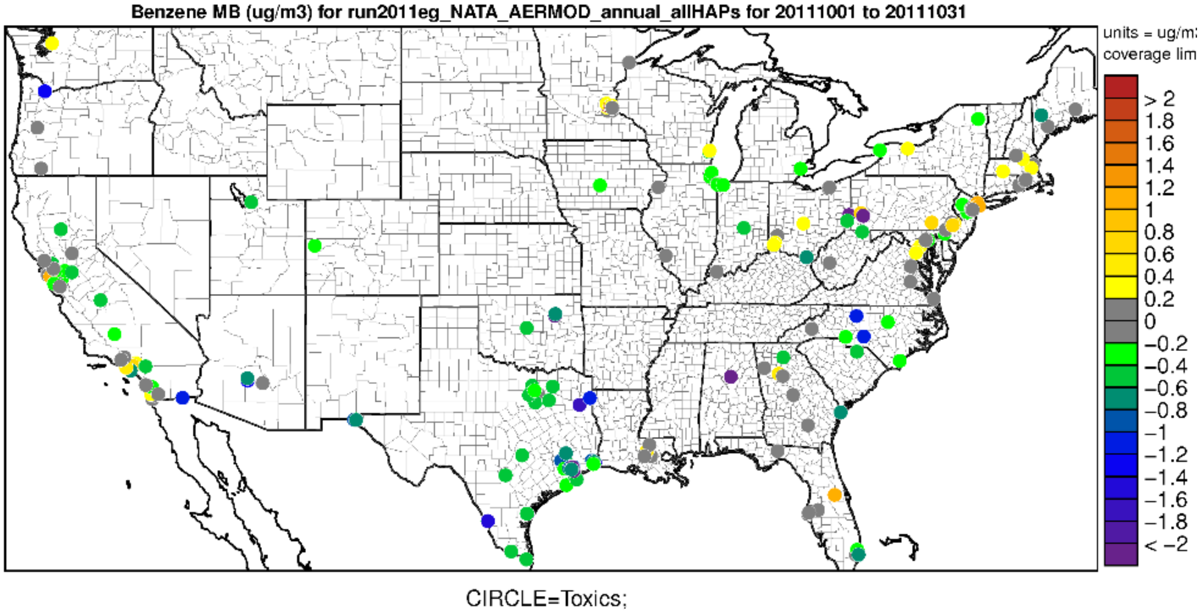


Exhibit F-24. Benzene: 2011 Mean Error (%) at Monitoring Sites in the HEM-3 Modeling Domain

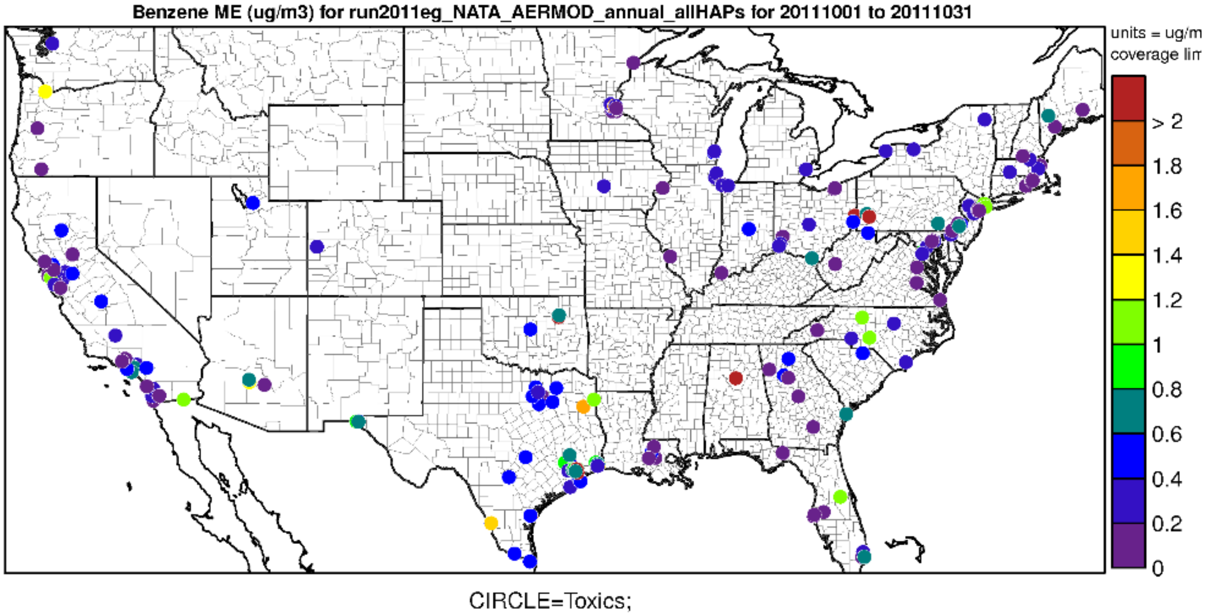


Exhibit F-25. 1,3-butadiene: 2011 Boxplots of Observed and Modeled Concentrations (top) and Modeled-Observed Bias Difference in Concentrations (bottom) for the Hybrid, CMAQ, and HEM-3 Models

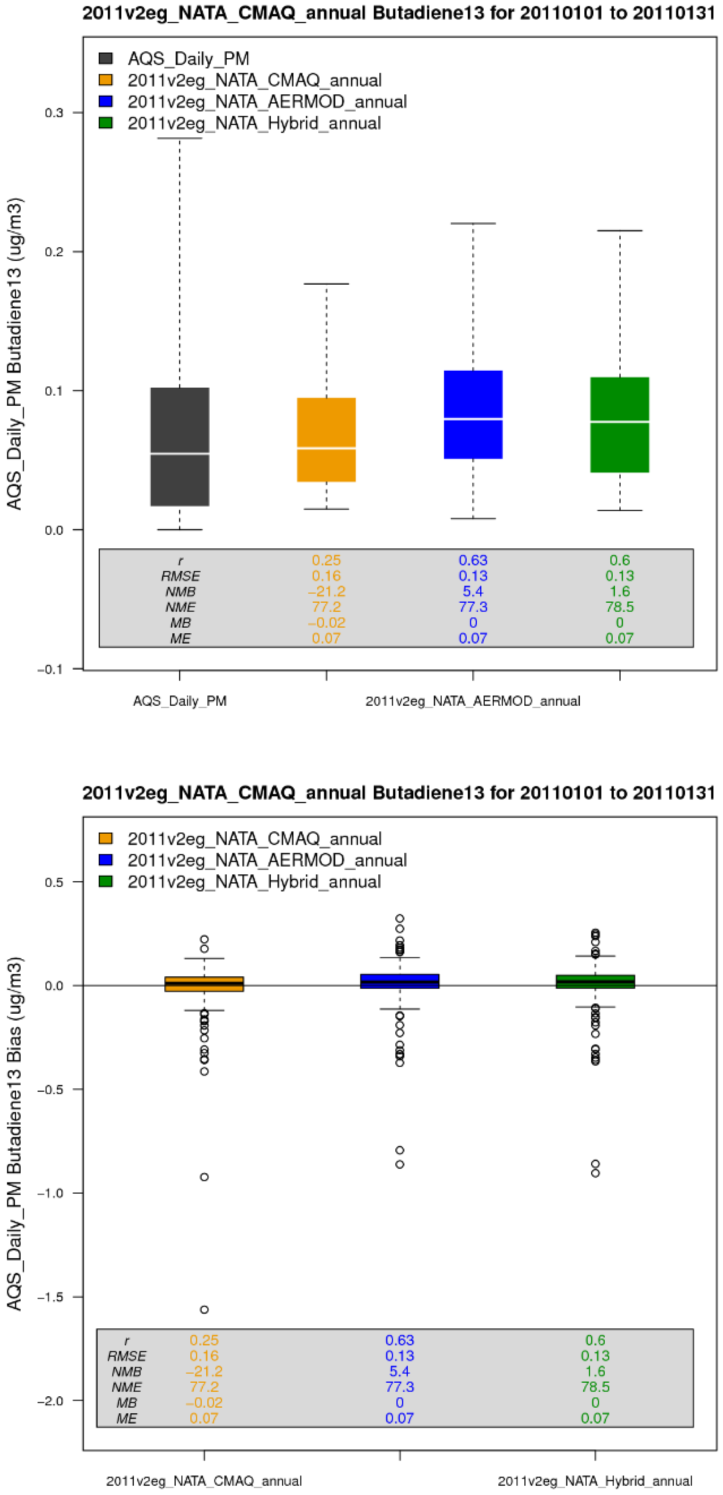


Exhibit F-26. 1,3-butadiene: 2011 Mean Bias (%) at Monitoring Sites in the Hybrid Modeling Domain

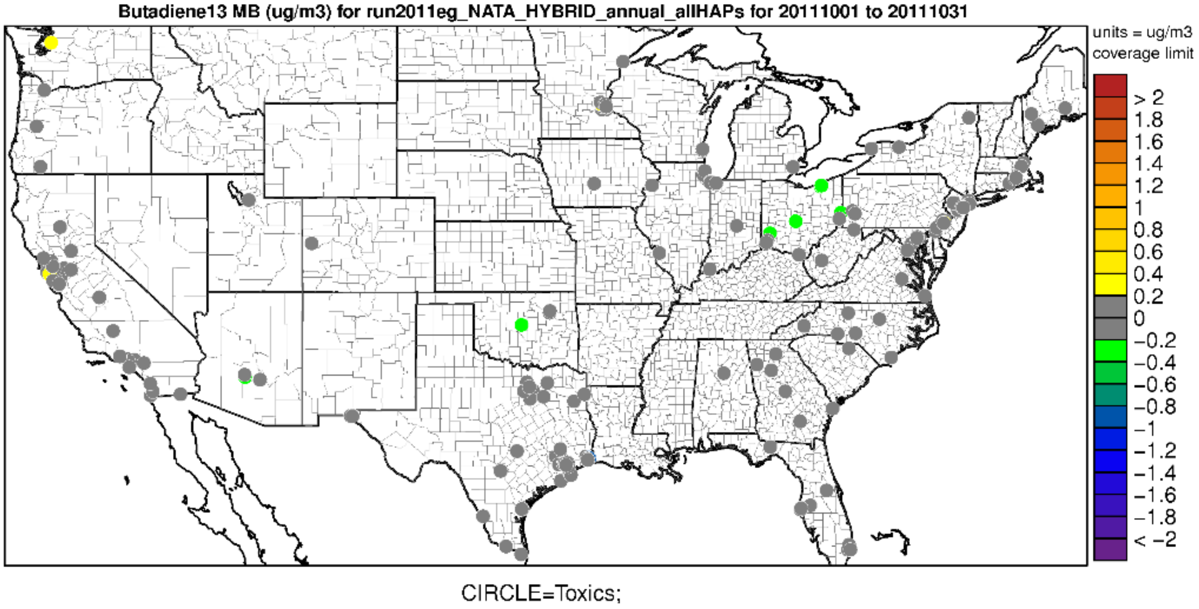


Exhibit F-27. 1,3-butadiene: 2011 Mean Error (%) at Monitoring Sites in the Hybrid Modeling Domain

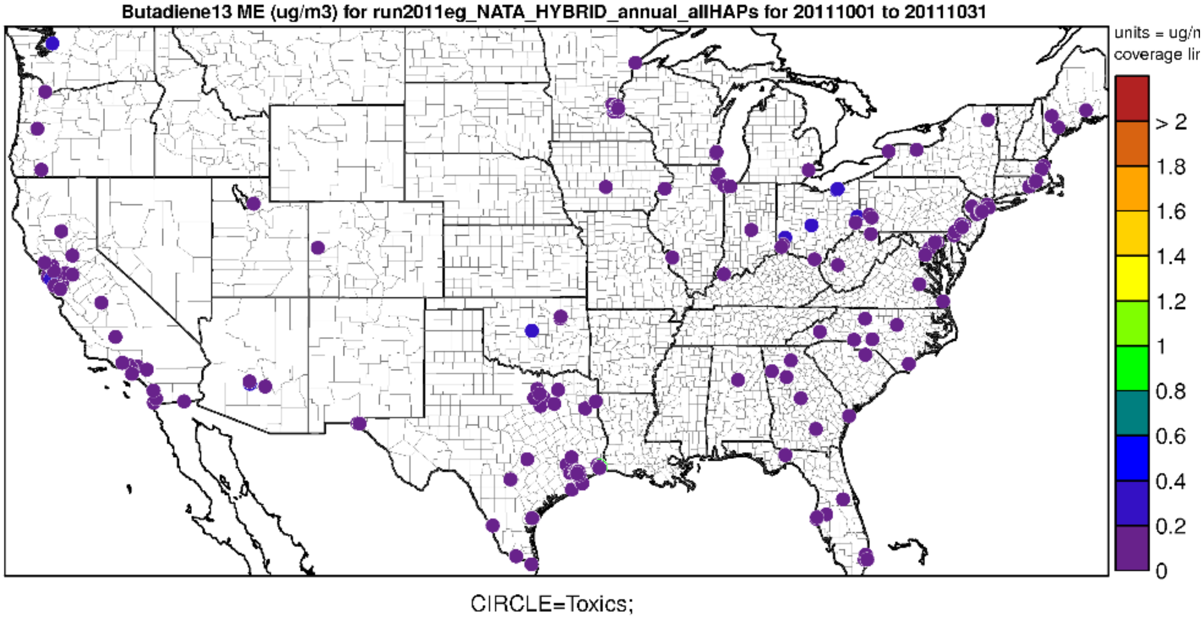


Exhibit F-28. 1,3-butadiene: 2011 Mean Bias (%) at Monitoring Sites in the CMAQ Modeling Domain

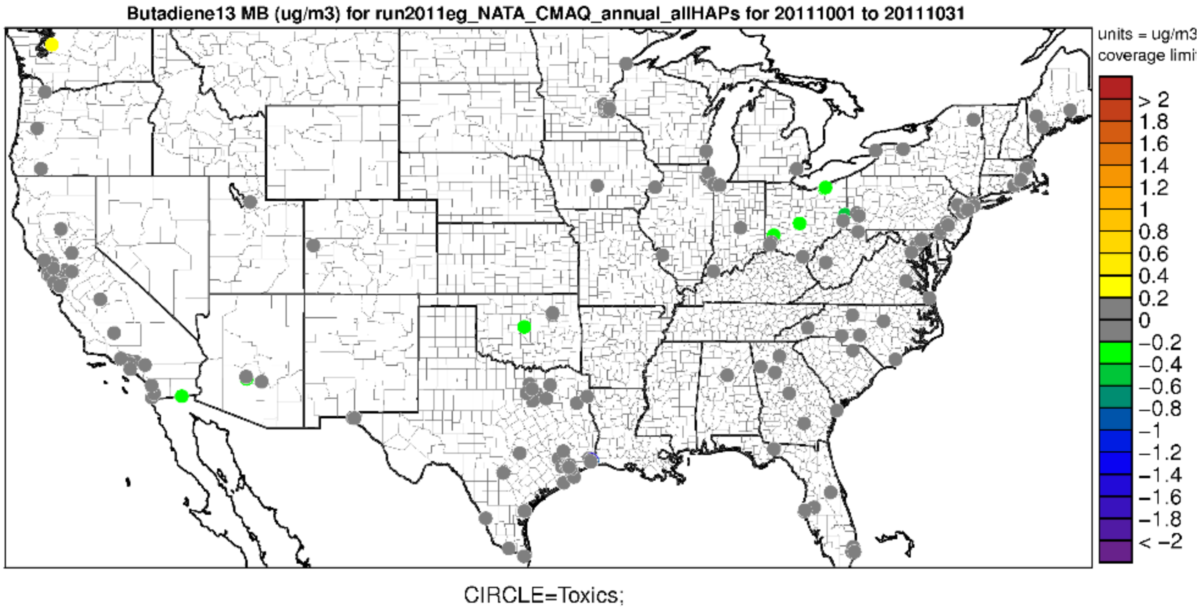


Exhibit F-29. 1,3-butadiene: 2011 Mean Error (%) at Monitoring Sites in the CMAQ Modeling Domain

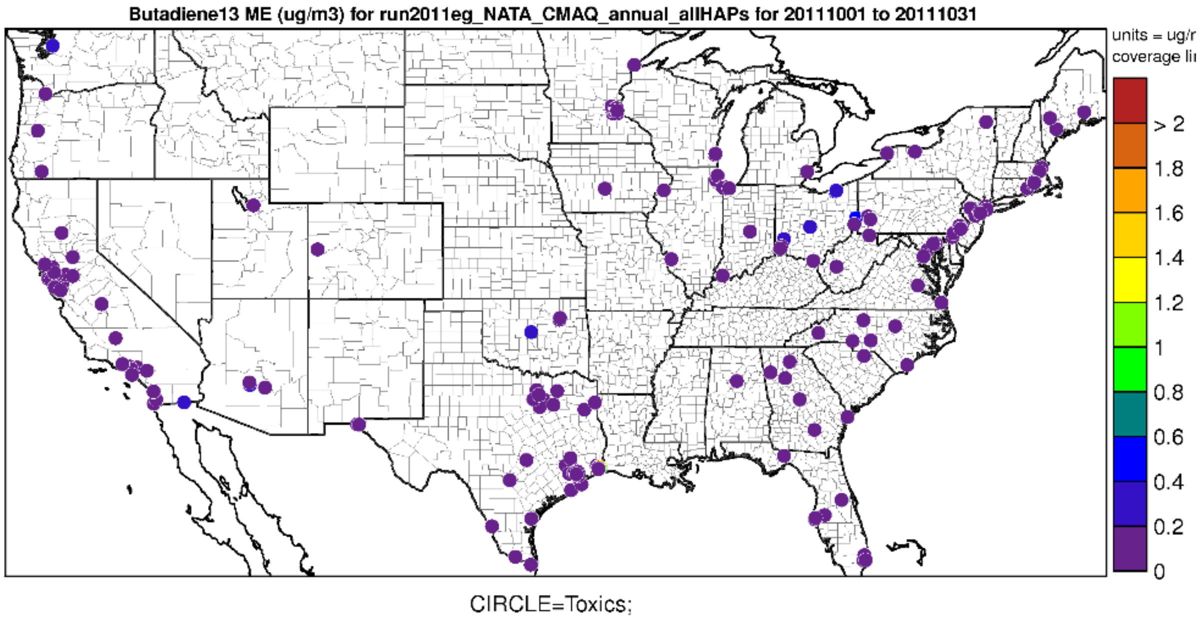


Exhibit F-30. 1,3-butadiene: 2011 Mean Bias (%) at Monitoring Sites in the HEM-3 Modeling Domain

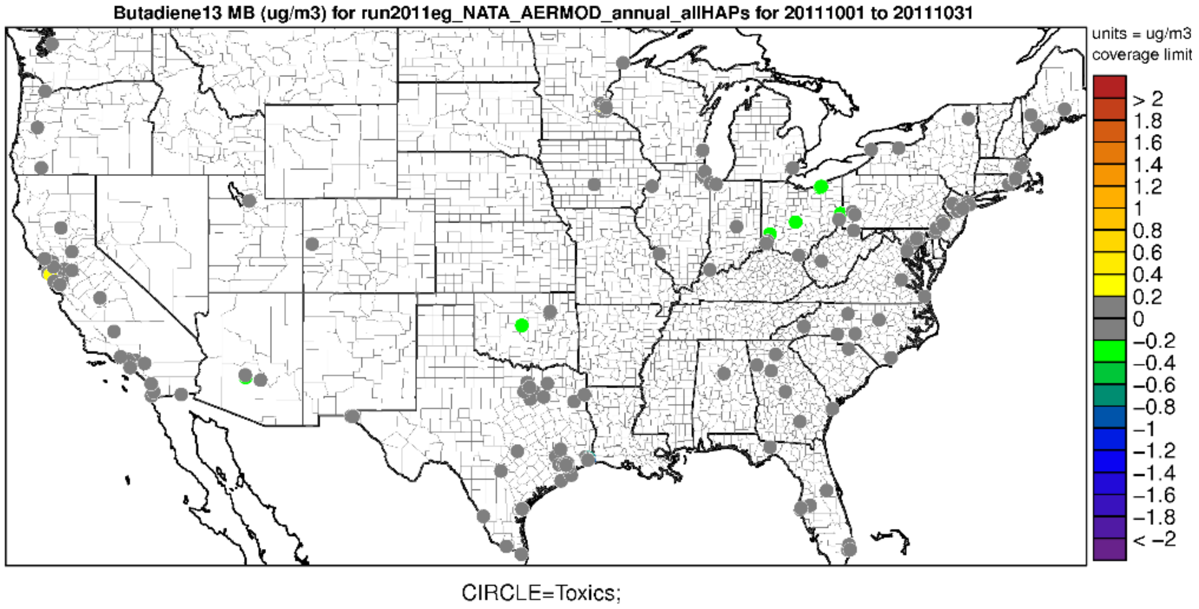


Exhibit F-31. 1,3-butadiene: 2011 Mean Error (%) at Monitoring Sites in the HEM-3 Modeling Domain

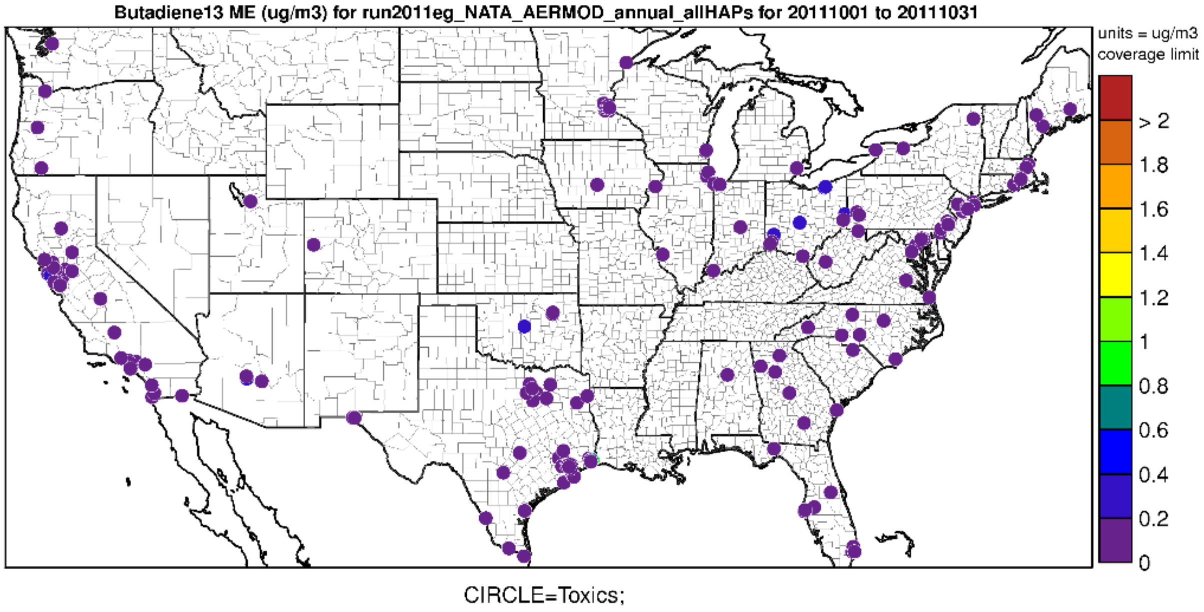
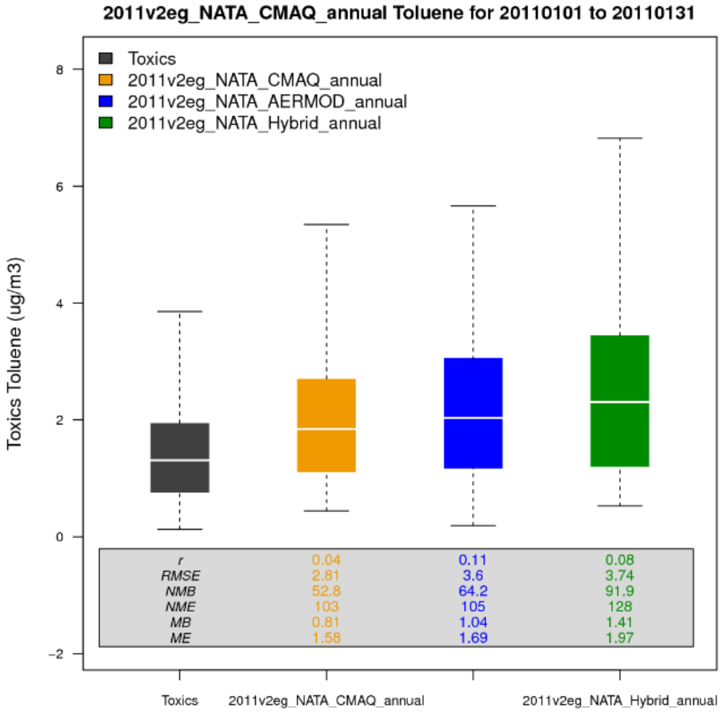


Exhibit F-32. Toluene: 2011 Boxplots of Observed and Modeled Concentrations (top) and Modeled-Observed Bias Difference in Concentrations (bottom) for the Hybrid, CMAQ, and HEM-3 Models



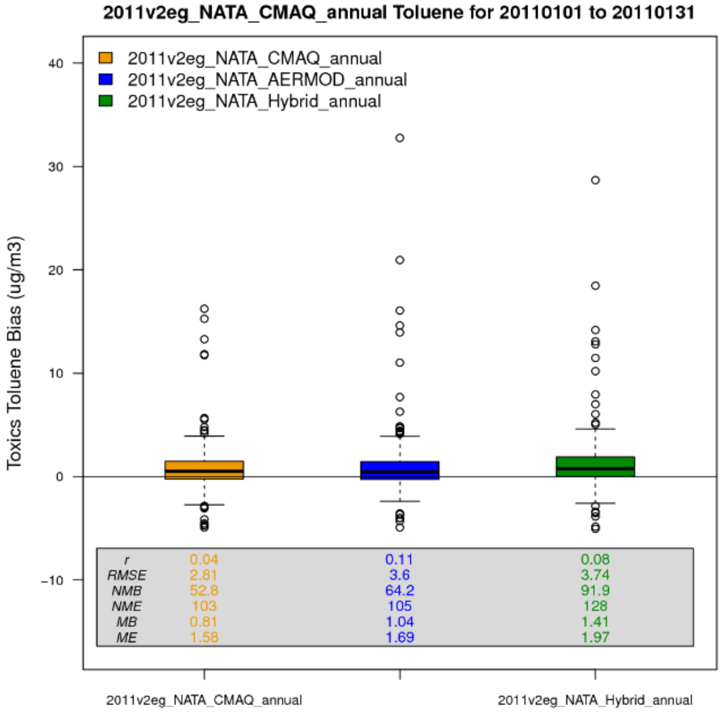


Exhibit F-33. Toluene: 2011 Mean Bias (%) at Monitoring Sites in the Hybrid Modeling Domain

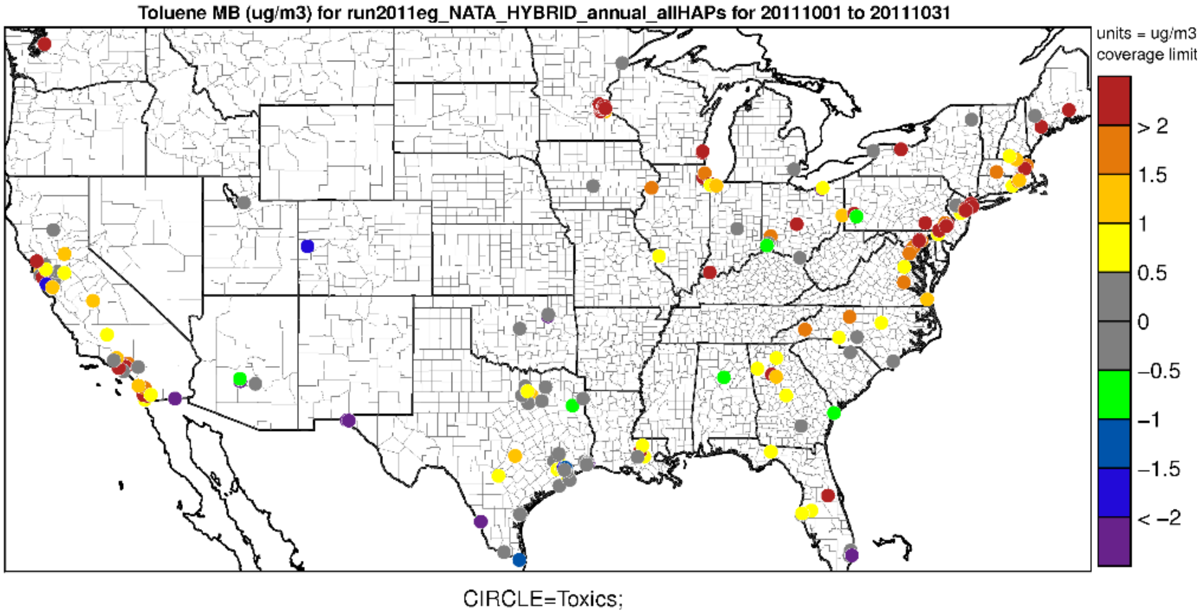


Exhibit F-34. Toluene: 2011 Mean Error (%) at Monitoring Sites in the Hybrid Modeling Domain

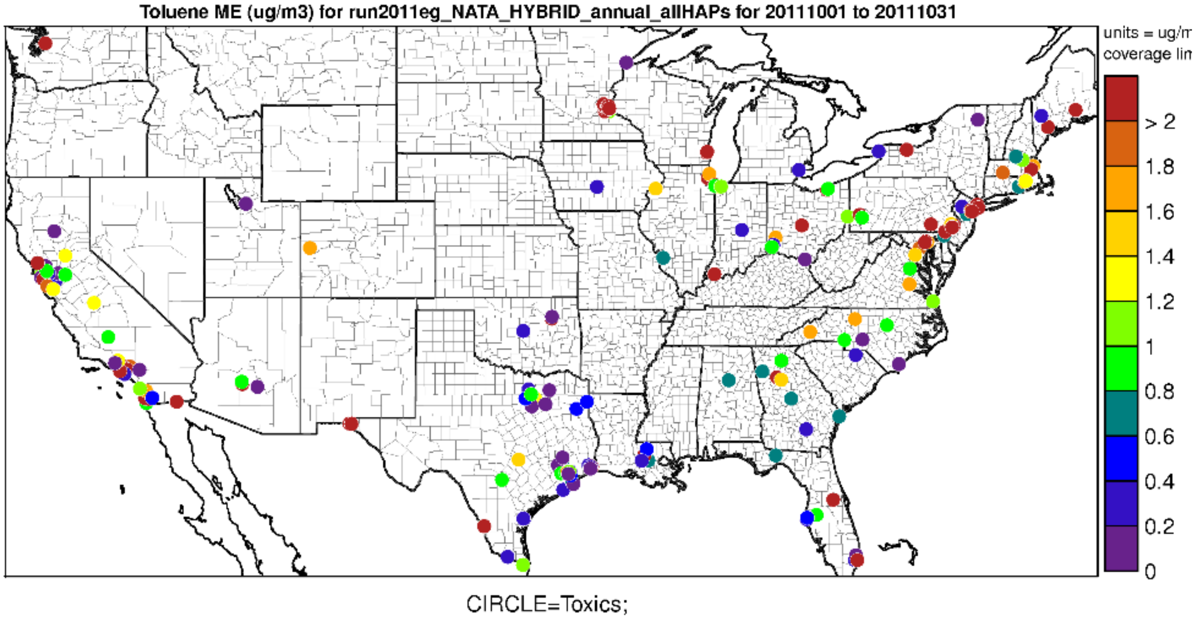


Exhibit F-35. Toluene: 2011 Mean Bias (%) at Monitoring Sites in the CMAQ Modeling Domain

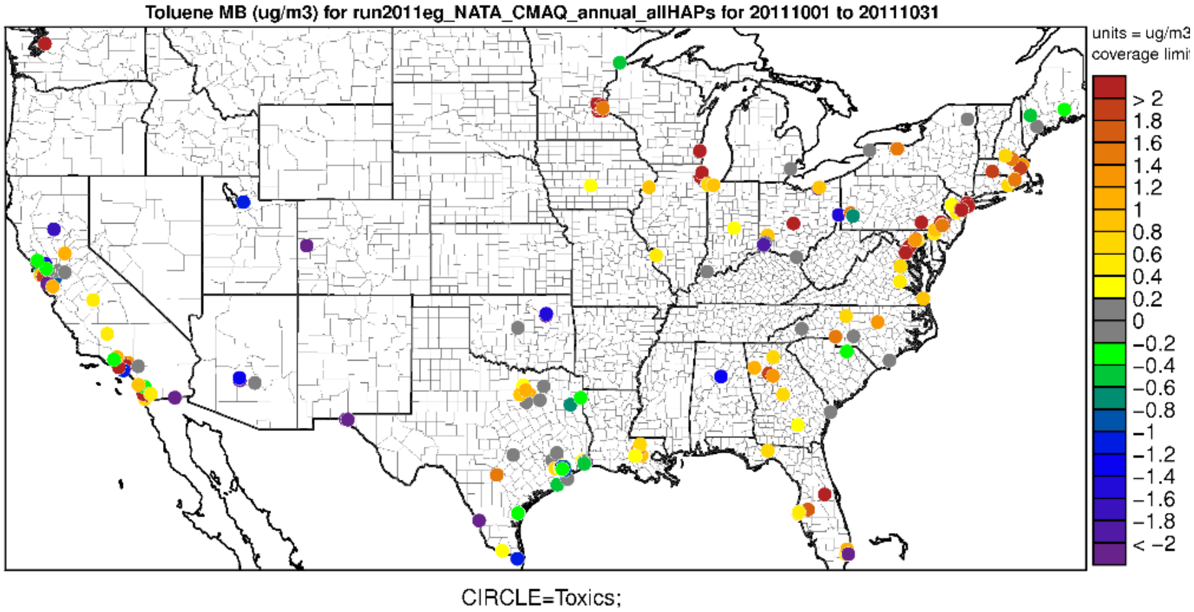


Exhibit F-36. Toluene: 2011 Mean Error (%) at Monitoring Sites in the CMAQ Modeling Domain

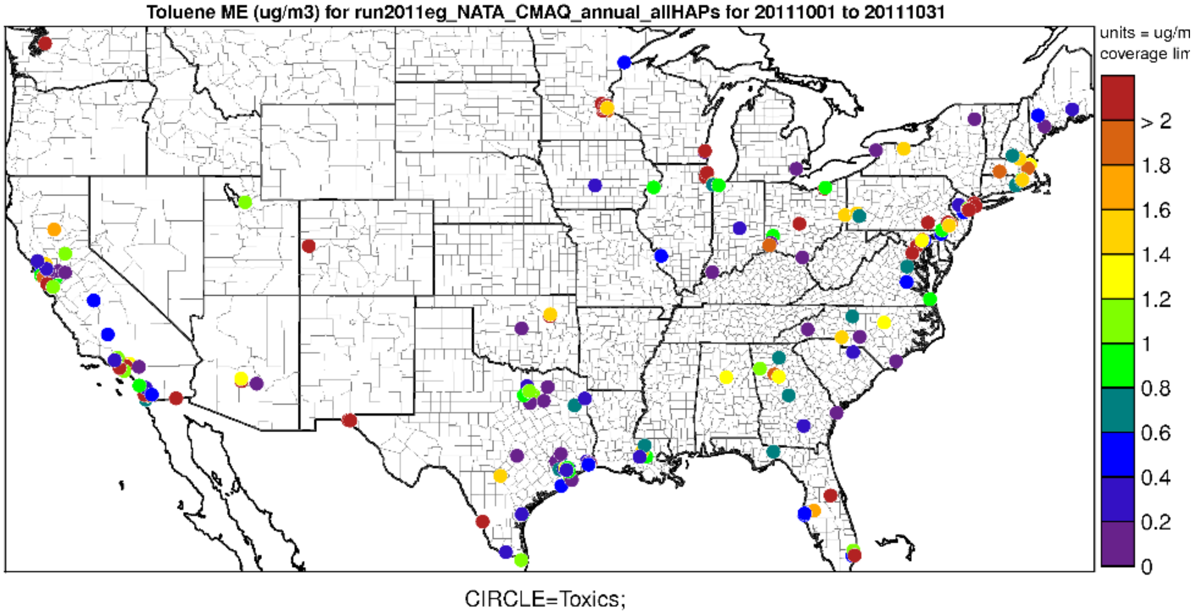


Exhibit F-37. Toluene: 2011 Mean Bias (%) at Monitoring Sites in the HEM-3 Modeling Domain

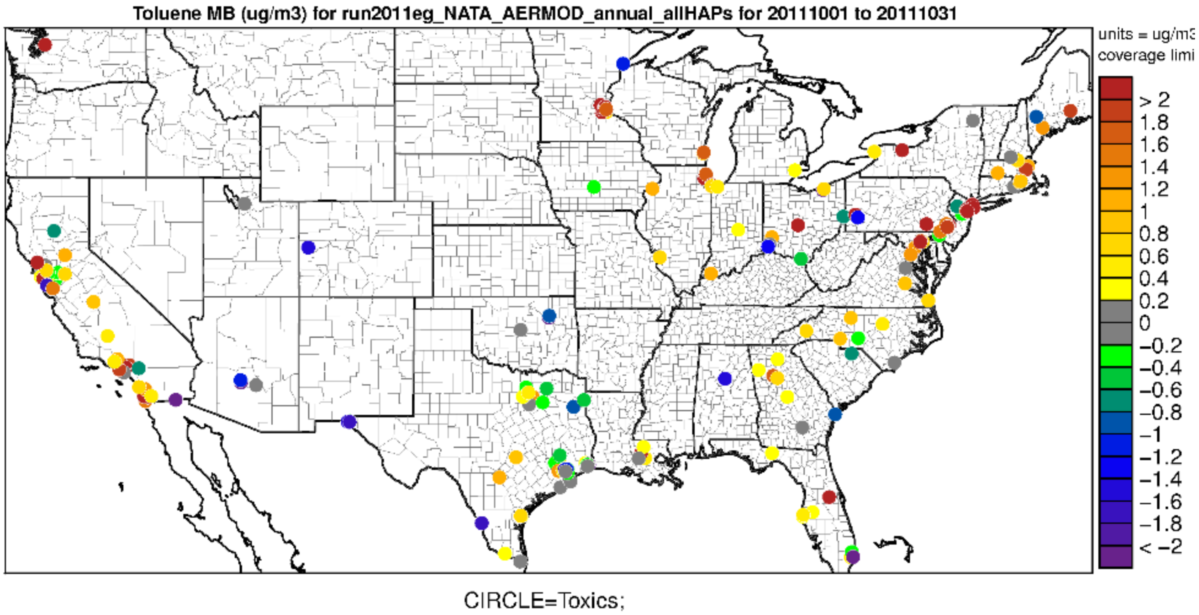


Exhibit F-38. Toluene: 2011 Mean Error (%) at Monitoring Sites in the HEM-3 Modeling Domain

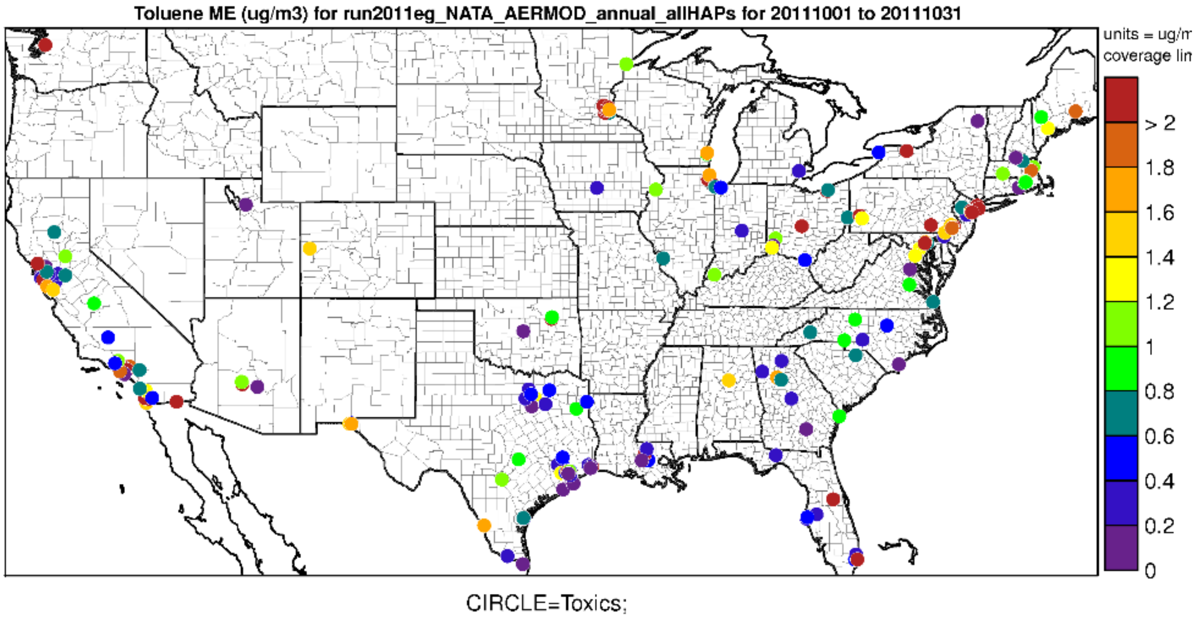


Exhibit F-39. Lead PM_{10}: 2011 Boxplots of Observed and Modeled Concentrations (top) and Modeled-Observed Bias Difference in Concentrations (bottom) for the Hybrid, CMAQ, and HEM-3 Models

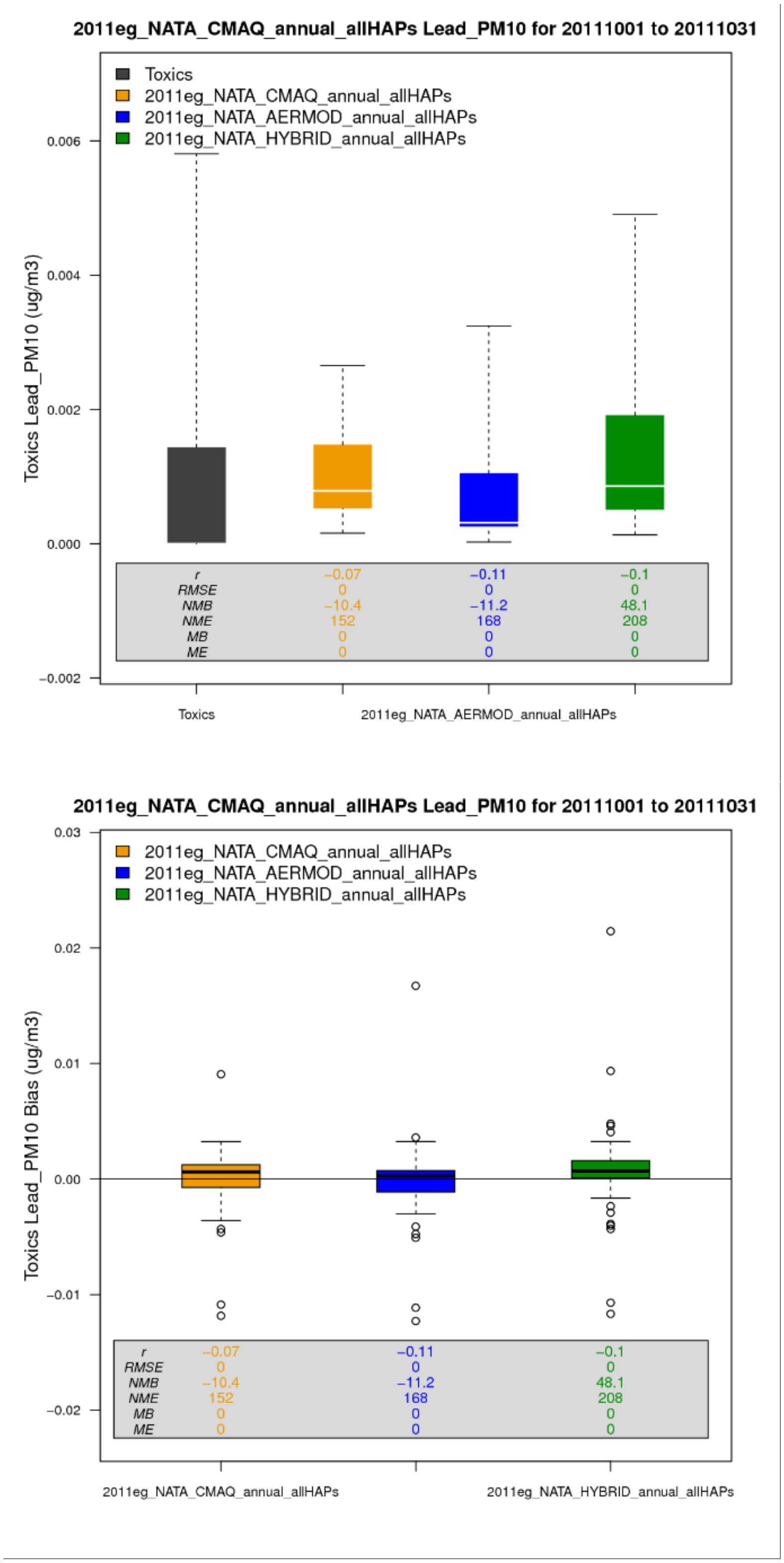


Exhibit F-40. Lead PM$_{10}$: 2011 Mean Bias (%) at Monitoring Sites in the Hybrid Modeling Domain

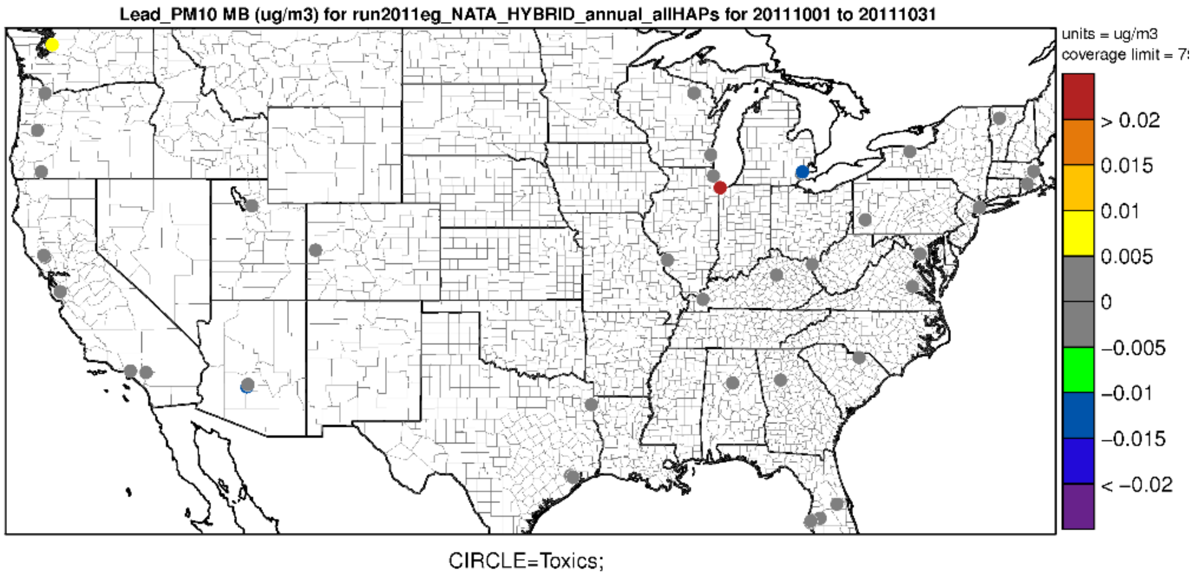


Exhibit F-41. Lead PM$_{10}$: 2011 Mean Error (%) at Monitoring Sites in the Hybrid Modeling Domain

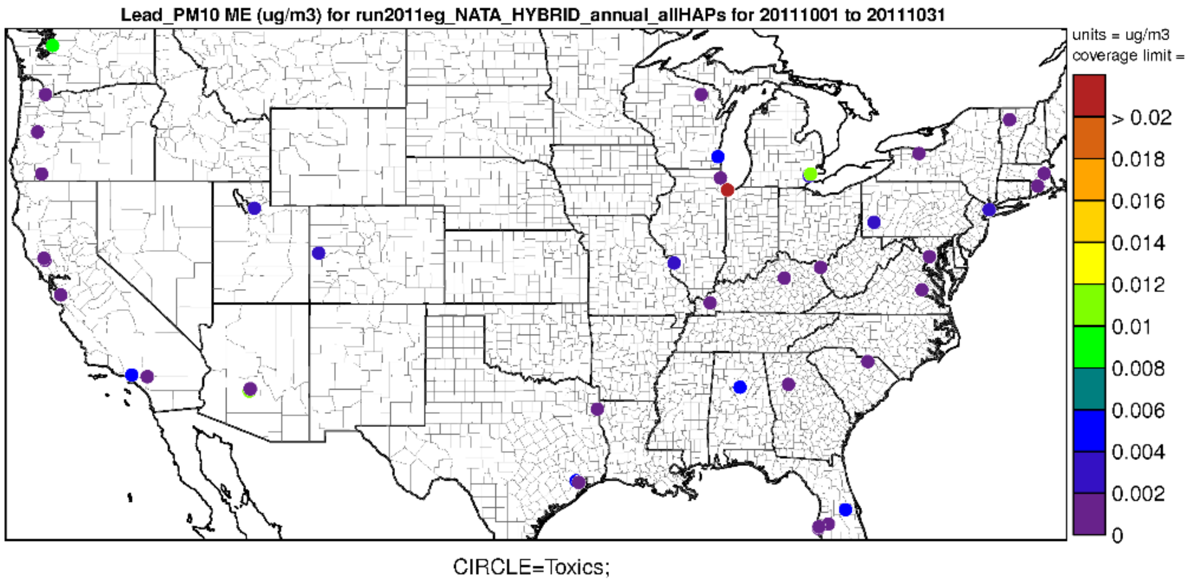


Exhibit F-42. Lead PM10: 2011 Mean Bias (%) at Monitoring Sites in the CMAQ Modeling Domain

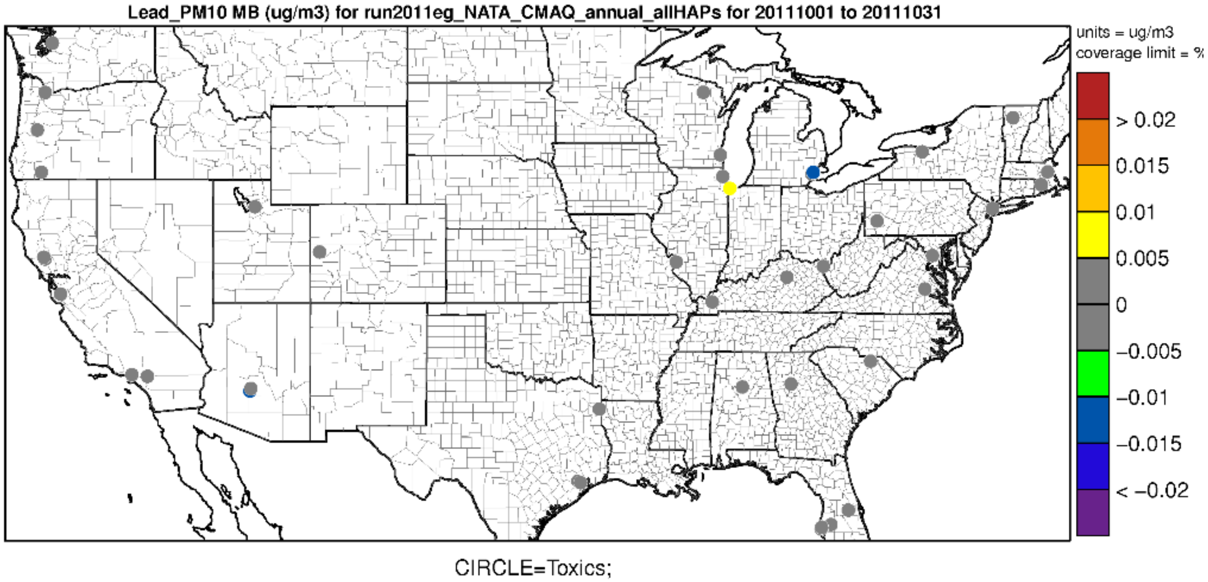


Exhibit F-43. Lead PM10: 2011 Mean Error (%) at Monitoring Sites in the CMAQ Modeling Domain

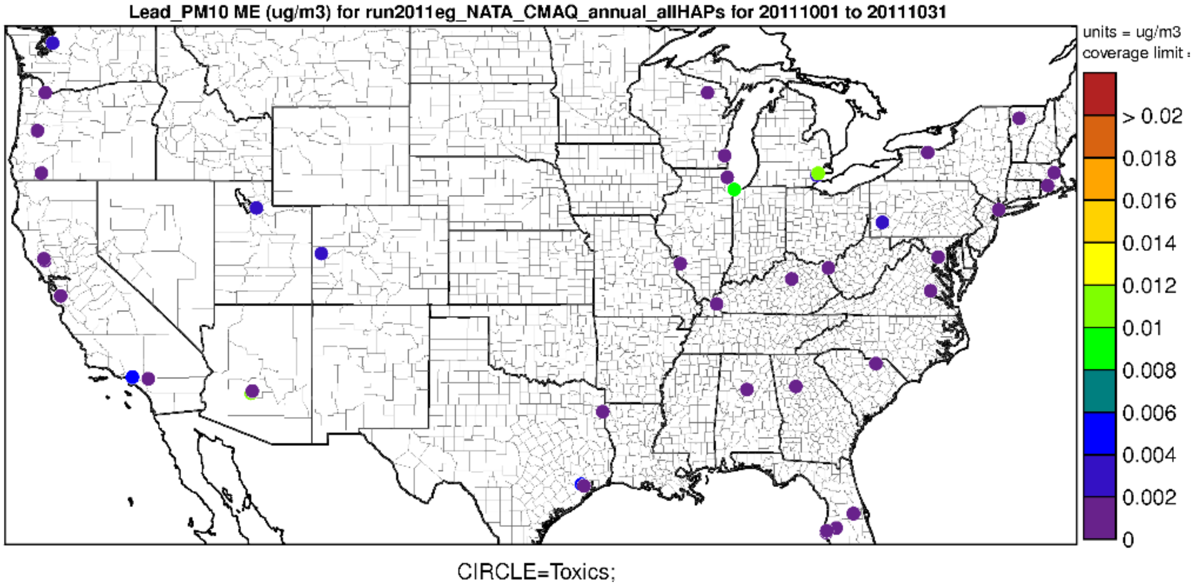


Exhibit F-44. Lead PM_{10}: 2011 Mean Bias (%) at Monitoring Sites in the HEM-3 Modeling Domain

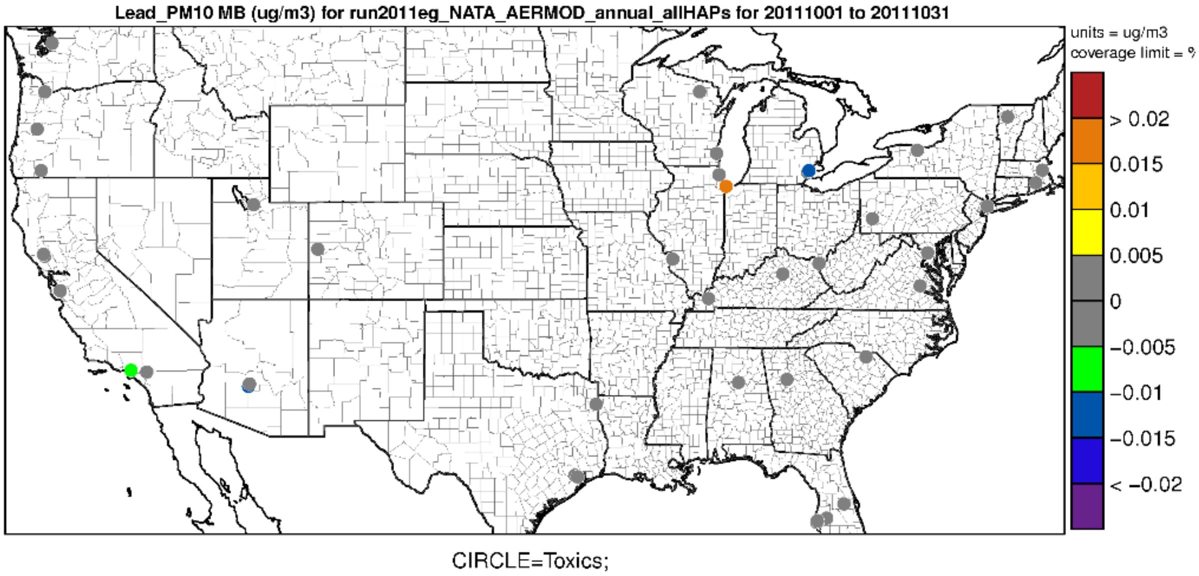


Exhibit F-45. Lead PM_{10}: 2011 Mean Error (%) at Monitoring Sites in the HEM-3 Modeling Domain

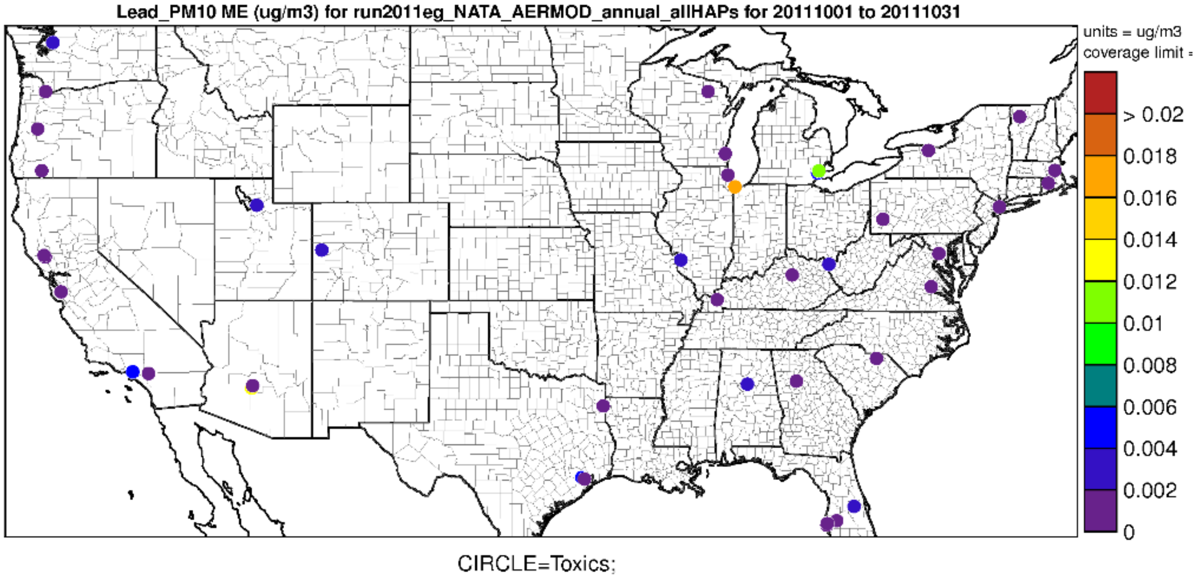
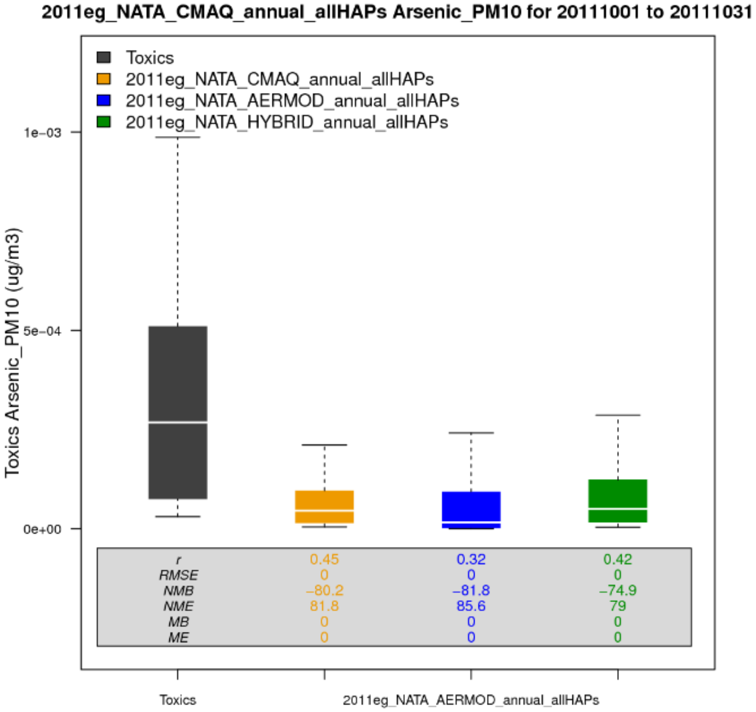


Exhibit F-46. Arsenic PM10: 2011 Boxplots of Observed and Modeled Concentrations (top) and Modeled-Observed Bias Difference in Concentrations (bottom) for the Hybrid, CMAQ, and HEM-3 Models



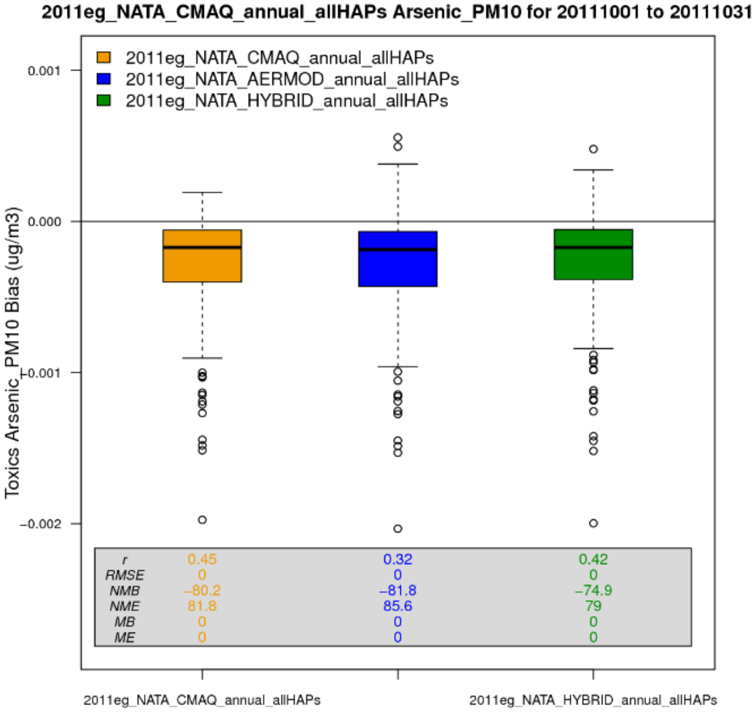


Exhibit F-47. Arsenic PM_{10}: 2011 Mean Bias (%) at Monitoring Sites in the Hybrid Modeling Domain

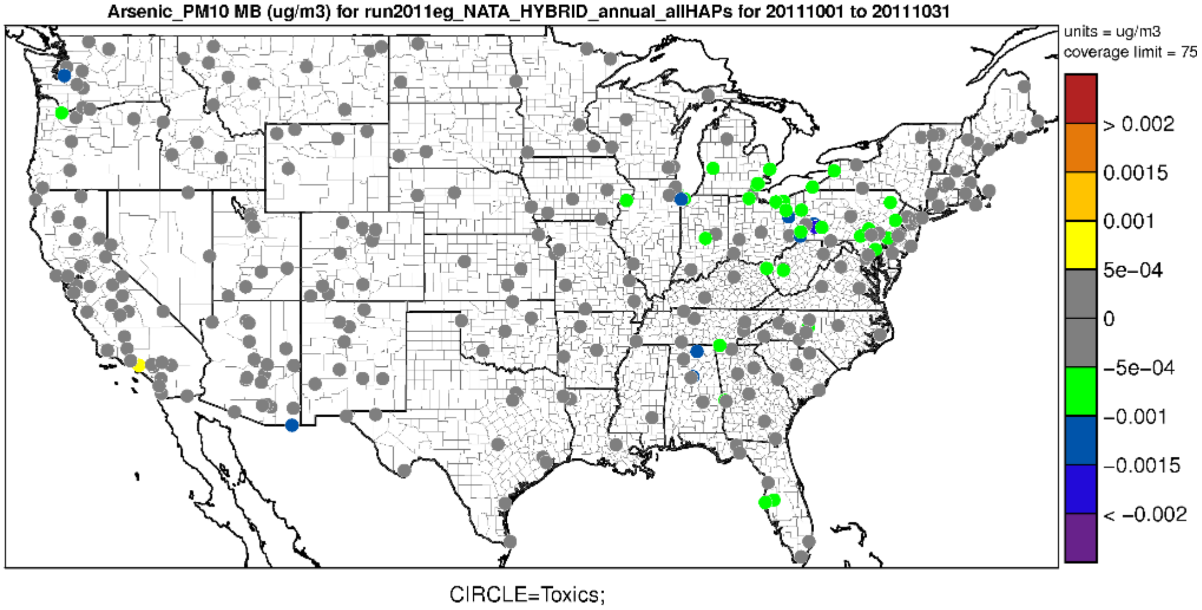


Exhibit F-48. Arsenic PM_{10}: 2011 Mean Error (%) at Monitoring Sites in the Hybrid Modeling Domain

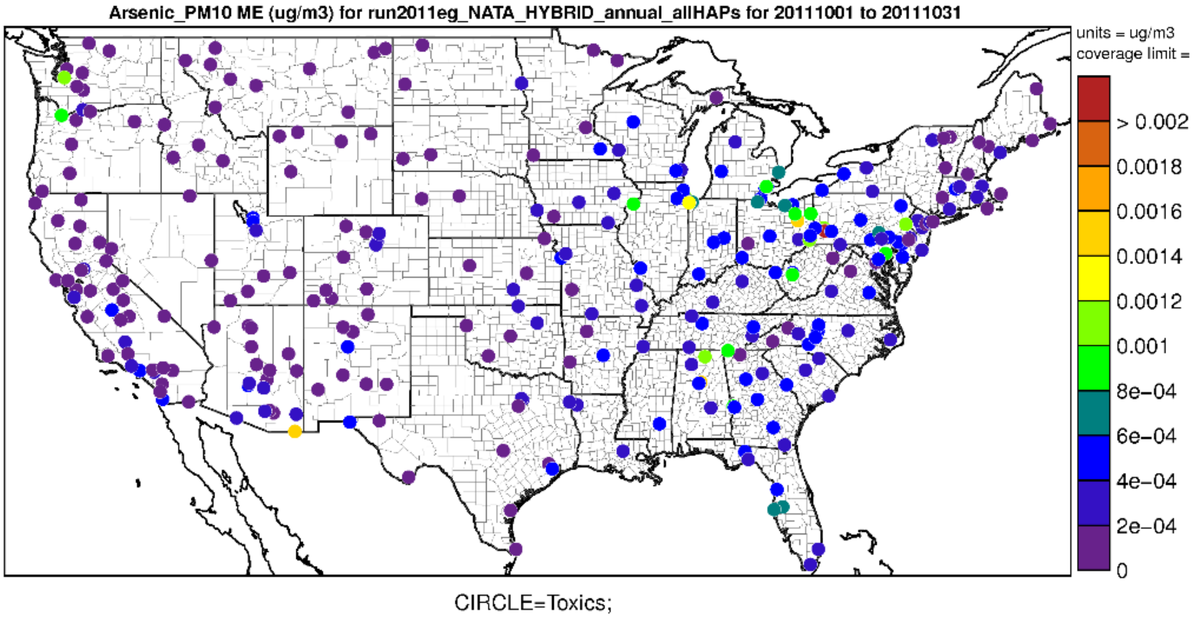
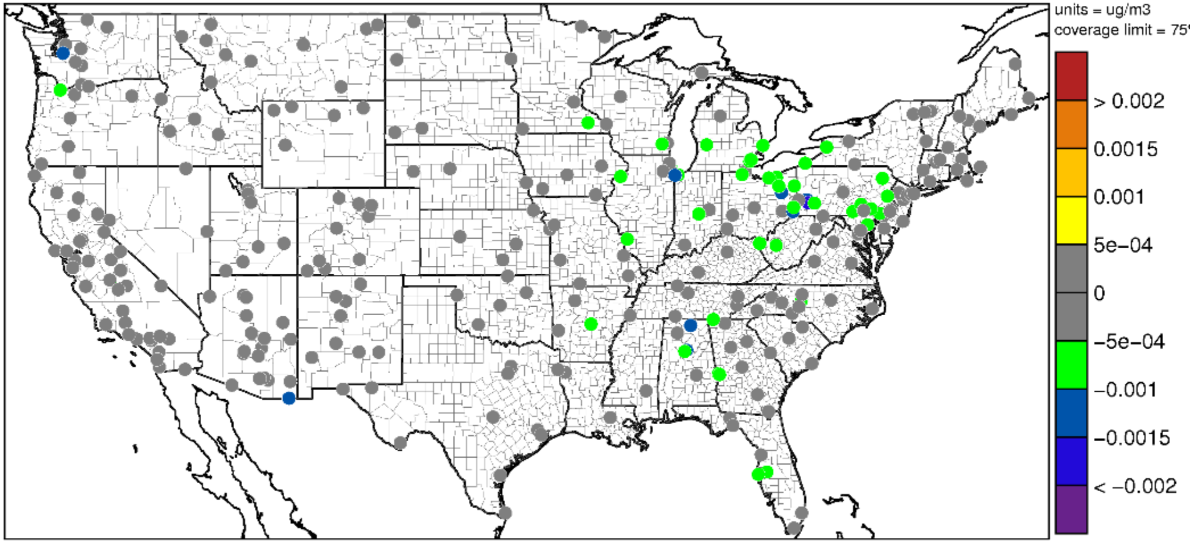


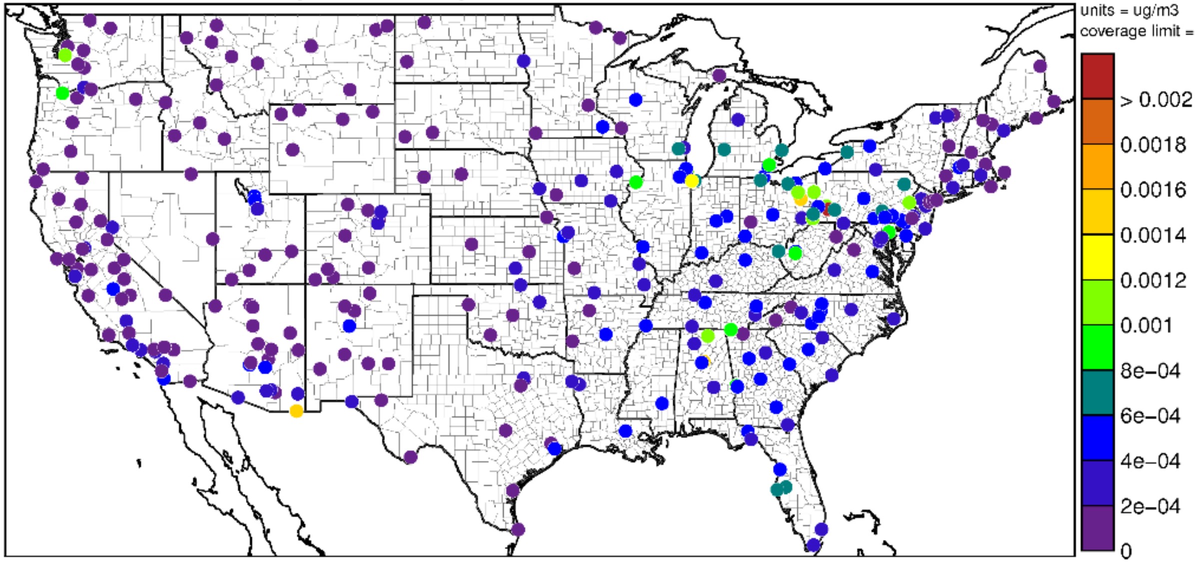
Exhibit F-49. Arsenic PM_{10}: 2011 Mean Bias (%) at Monitoring Sites in the CMAQ Modeling Domain



Arsenic_PM10 MB (ug/m3) for run2011eg_NATA_CMAQ_annual_allHAPs for 20111001 to 20111031

CIRCLE=Toxics;

Exhibit F-50. Arsenic PM_{10}: 2011 Mean Error (%) at Monitoring Sites in the CMAQ Modeling Domain



Arsenic_PM10 ME (ug/m3) for run2011eg_NATA_CMAQ_annual_allHAPs for 20111001 to 20111031

CIRCLE=Toxics;

Exhibit F-51. Arsenic PM$_{10}$: 2011 Mean Bias (%) at Monitoring Sites in the HEM-3 Modeling Domain

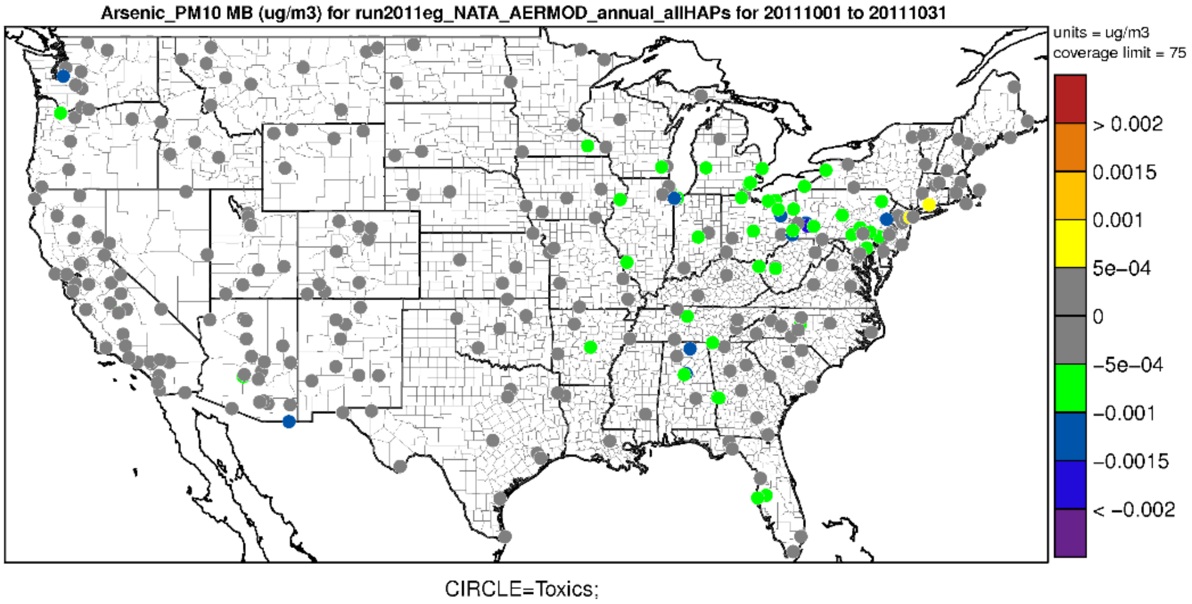


Exhibit F-52. Arsenic PM$_{10}$: 2011 Mean Error (%) at Monitoring Sites in the HEM-3 Modeling Domain

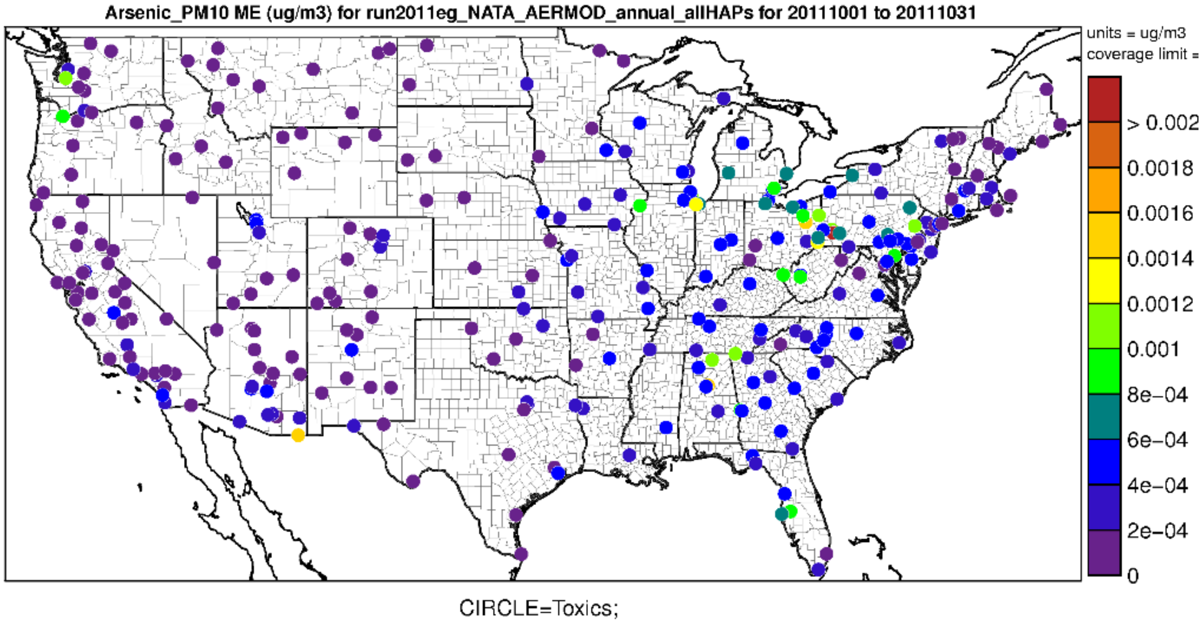
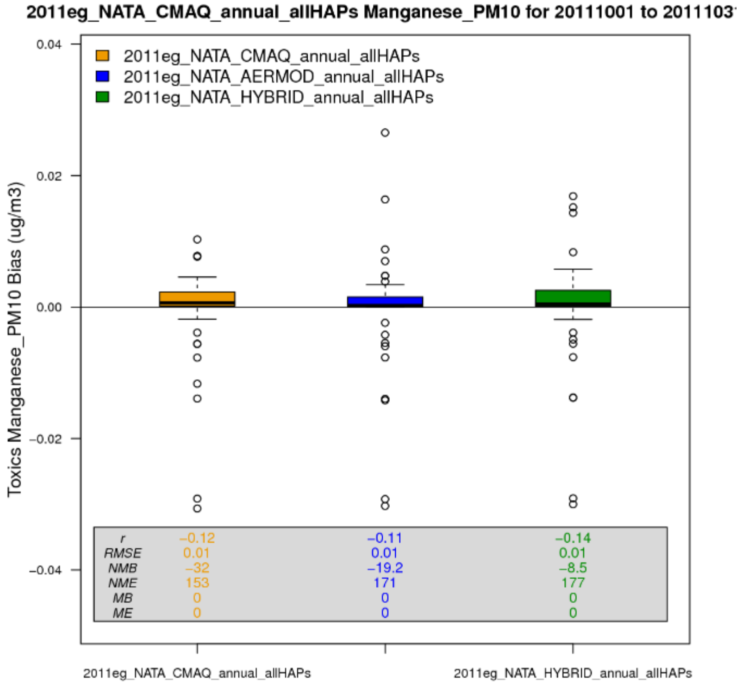
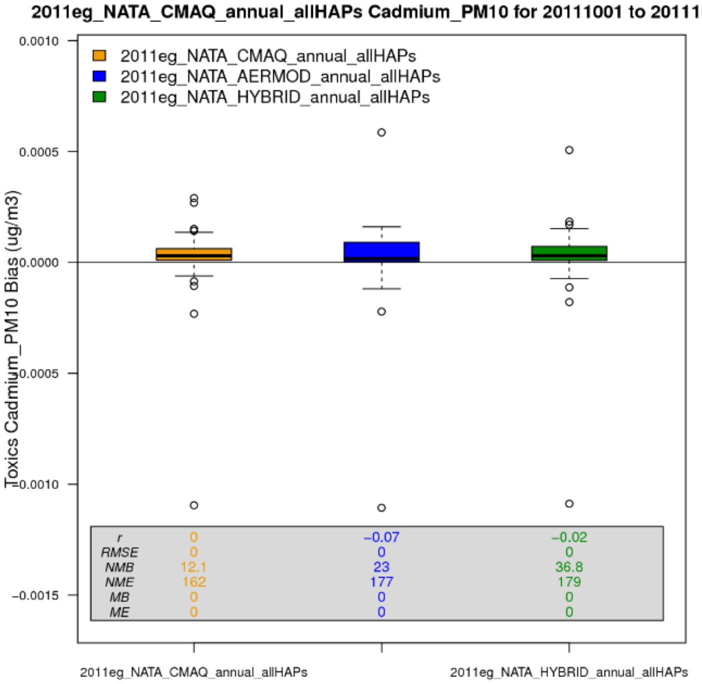


Exhibit F-53. Manganese PM$_{10}$ (a), Cadmium PM$_{10}$ (b), Nickel PM$_{10}$ (c), Naphthalene (d), Dichlorproprene [1,3-dichloropropene] (e), Dichlorobenzene [1,4-dichlorobenzene] (f), Tetrachloroethylene (g), Trichloroethylene (h), Methylene Chloride (i), 1,2-dichloroethane [ethylene dichloride] (j), Chlorine (k), Chloroform (l), Acrylonitrile (m), Propdichloride [propylene dichloride] (n), and Xylenes (o): 2011 Boxplots of Modeled-Observed Bias Difference in Concentrations for the Hybrid, CMAQ, and HEM-3 Models

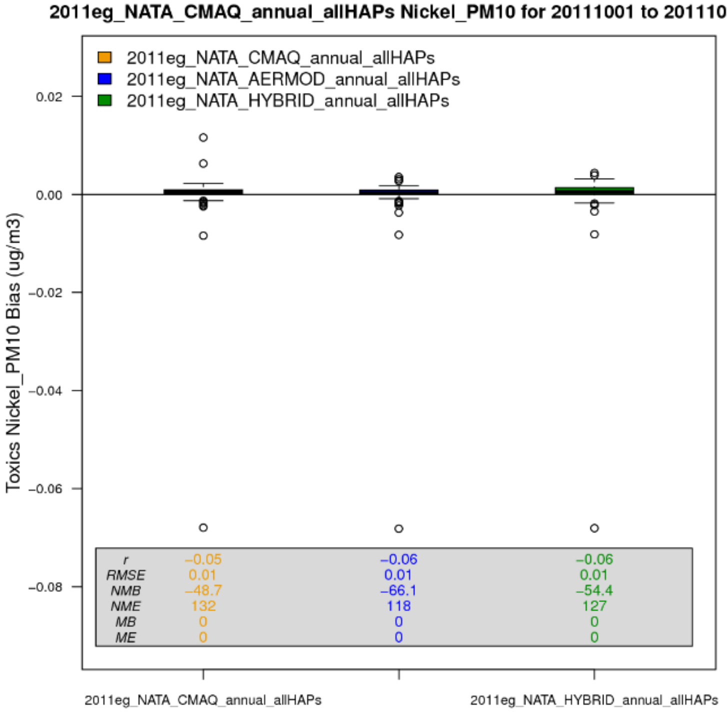
(a) Manganese PM$_{10}$:



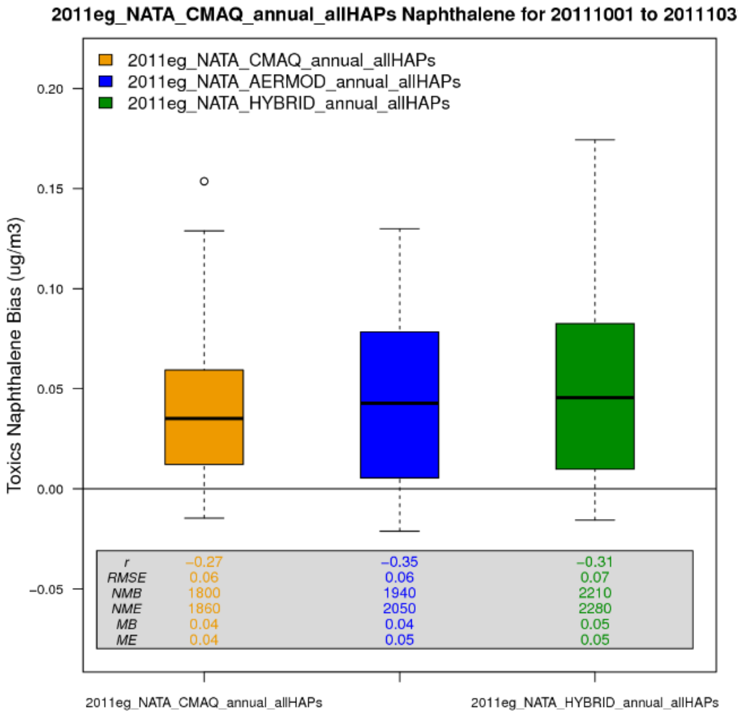
(b) Cadmium PM_{10}:



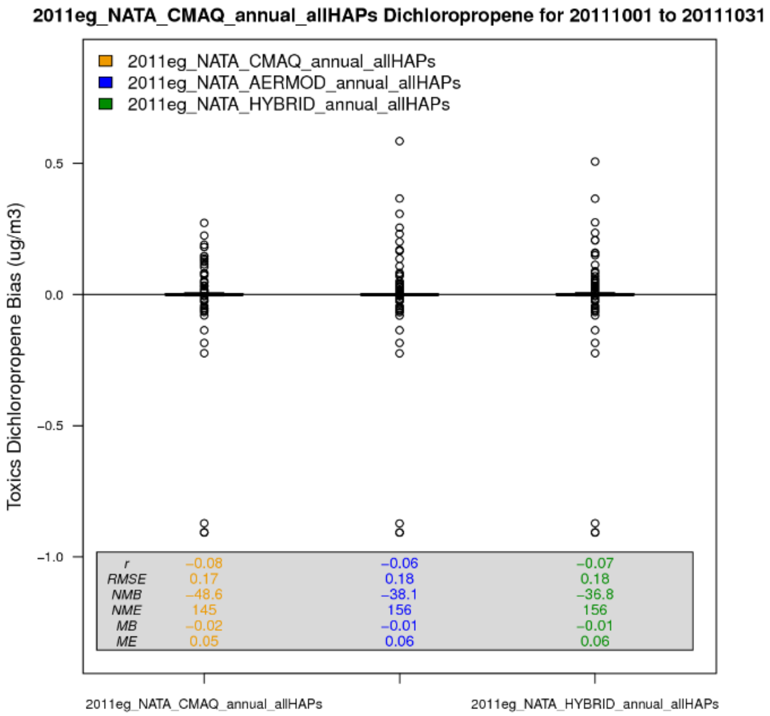
(c) Nickel PM_{10}:



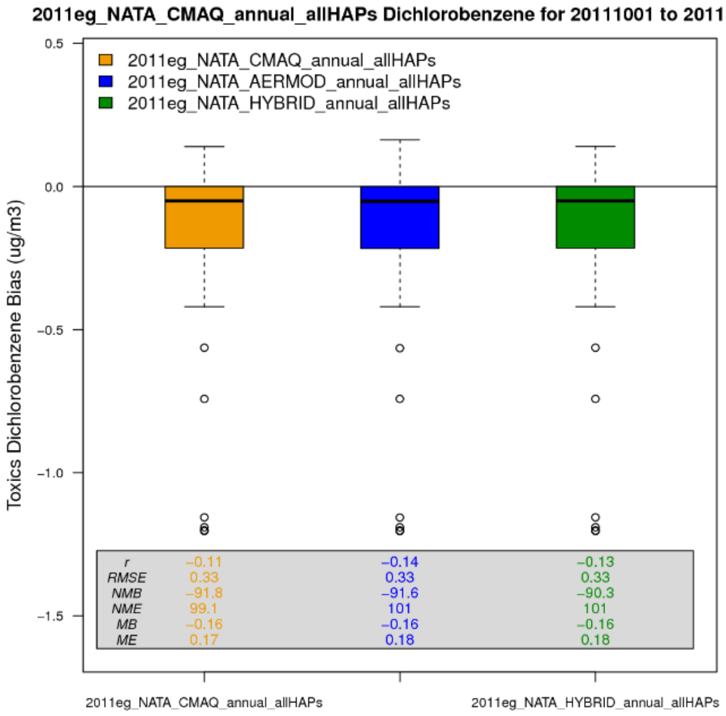
(d) Naphthalene:



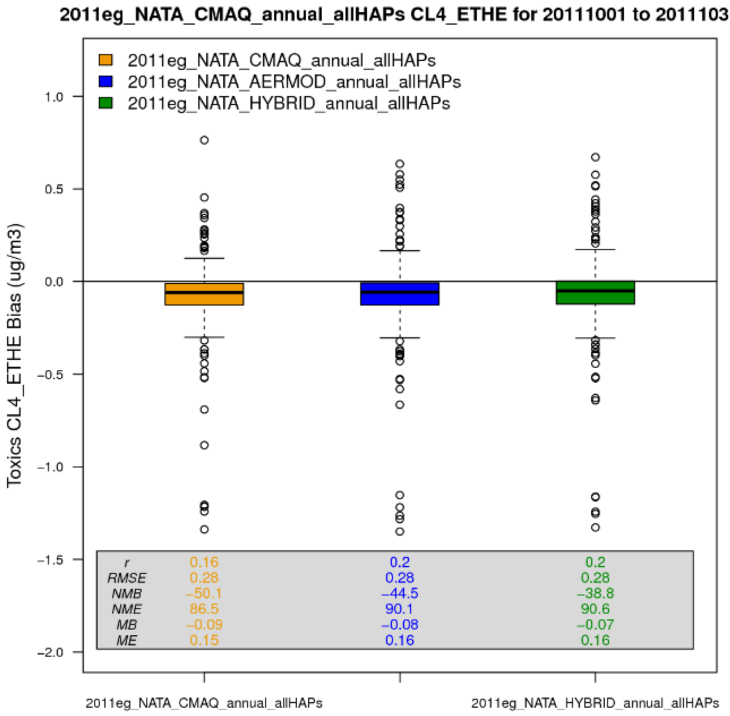
(e) Dichloropropene:



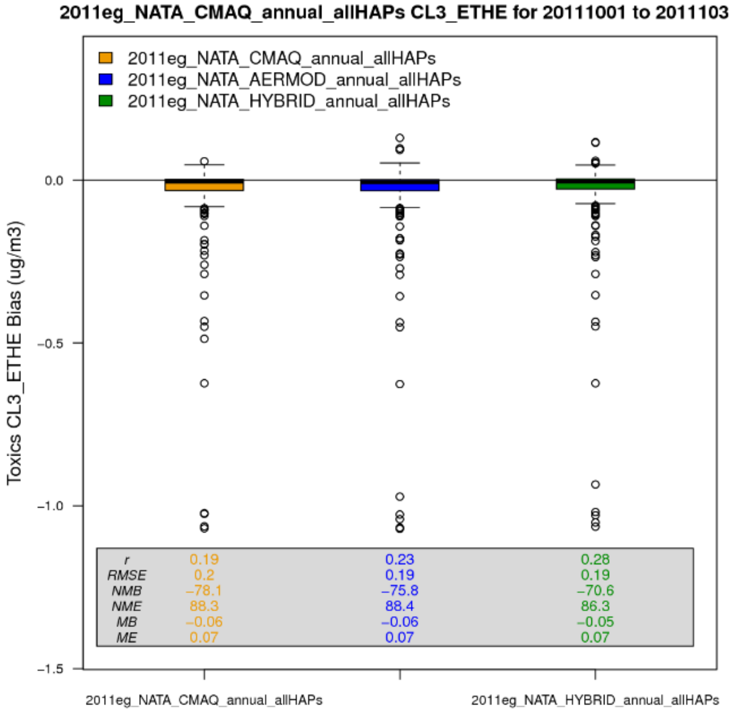
(f) Dichlorobenzene:



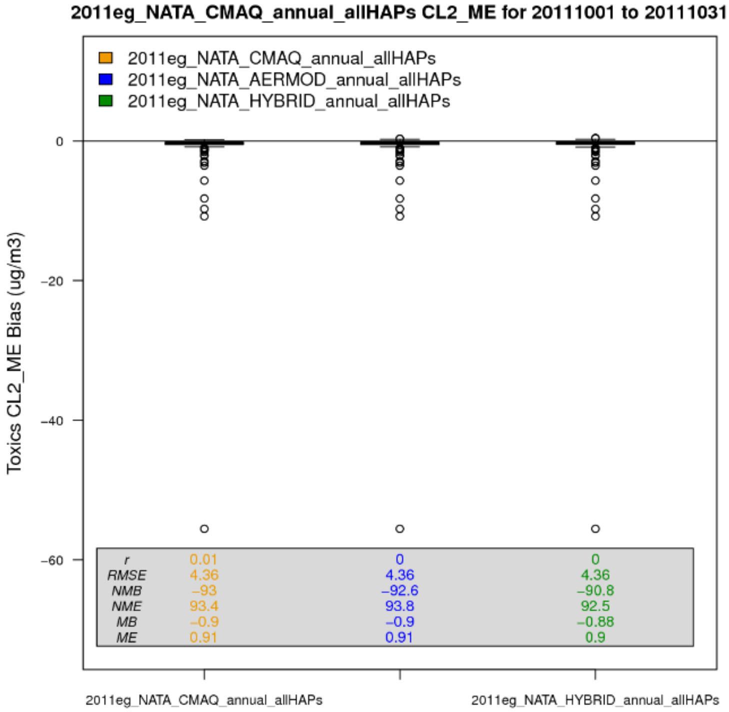
(g) Tetrachloroethylene:



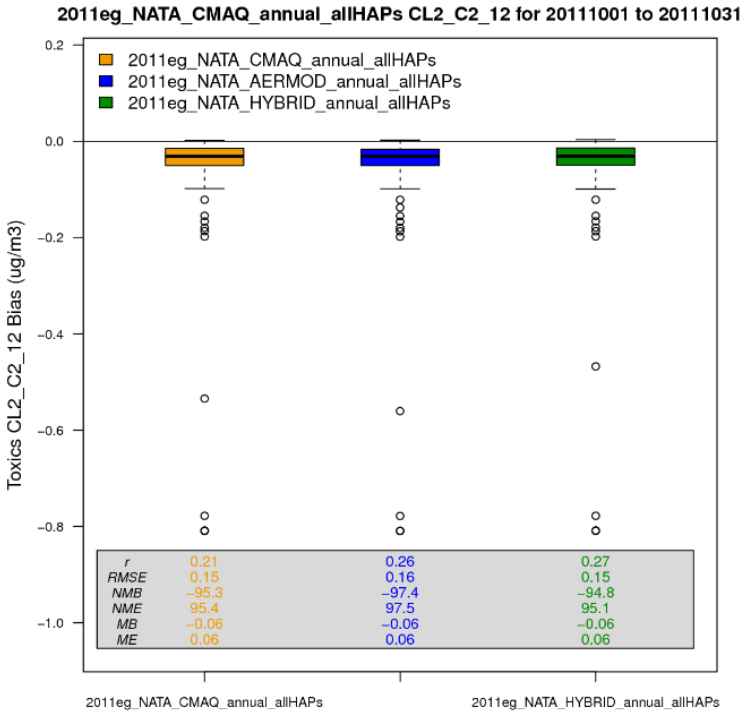
(h) Trichloroethylene:



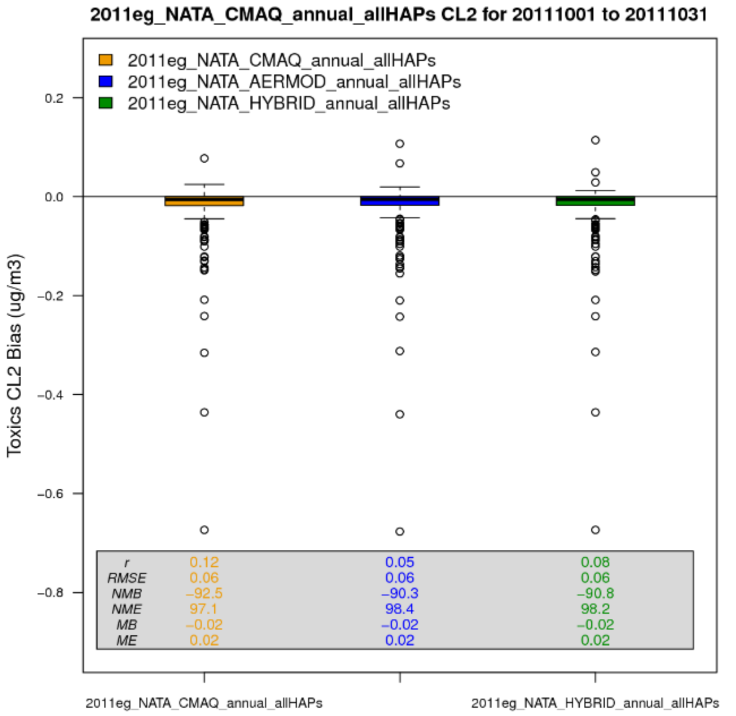
(i) Methylene Chloride:



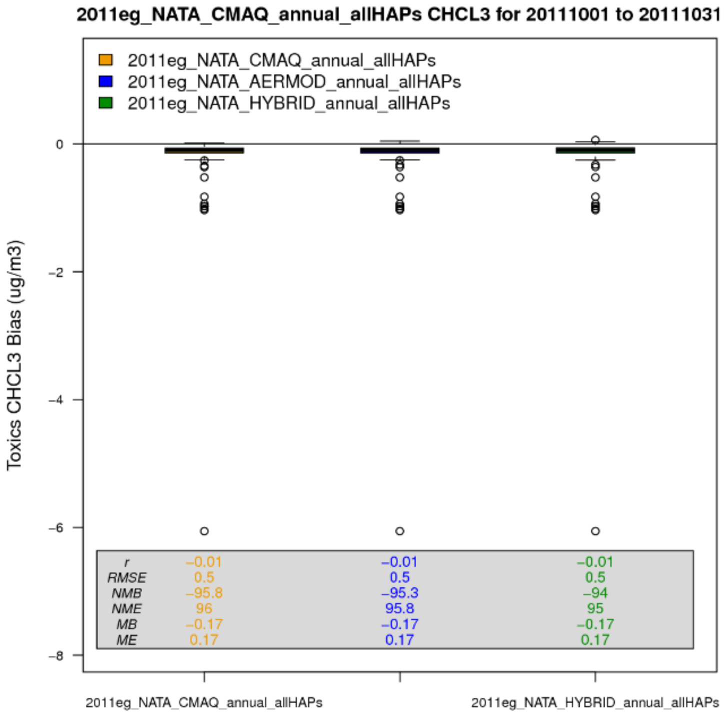
(j) 1,2-dichloroethane:



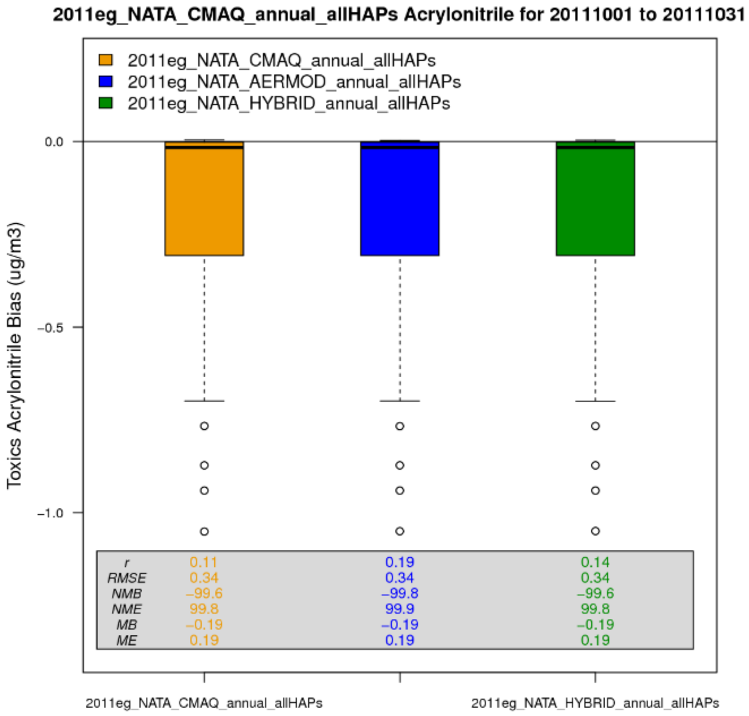
(k) Chlorine:



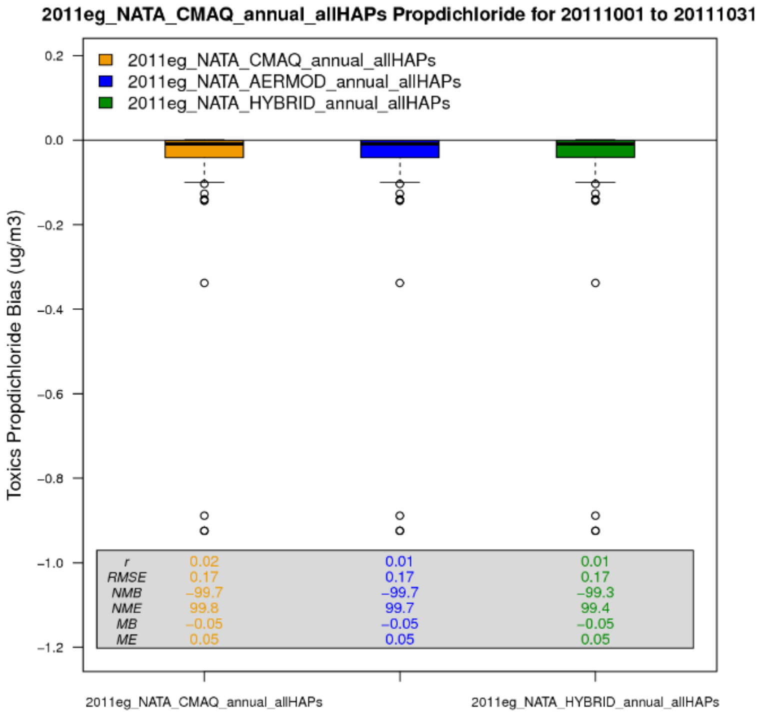
(l) Chloroform:



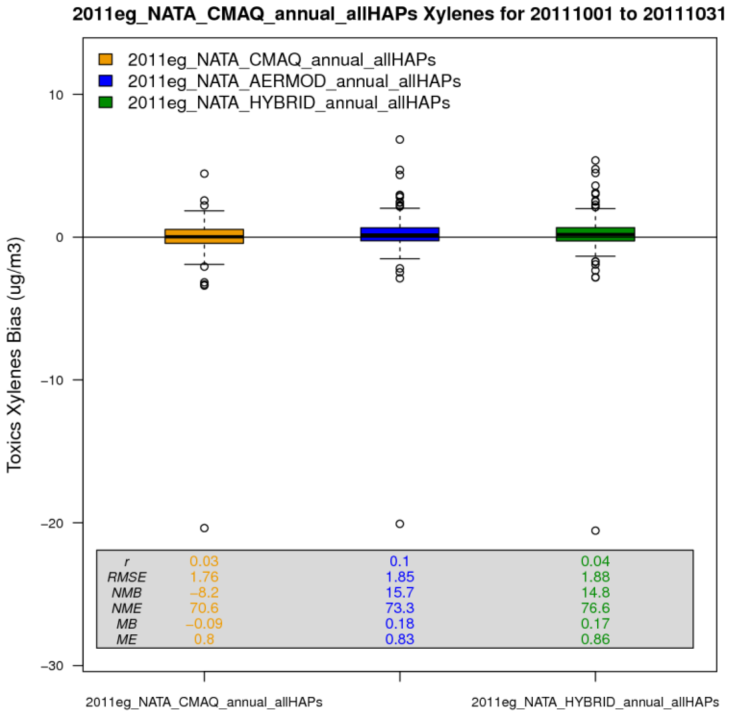
(m) Acrylonitrile:



(n) Propdichloride:



(o) Xylenes:



F.5 Non-hybrid Evaluation

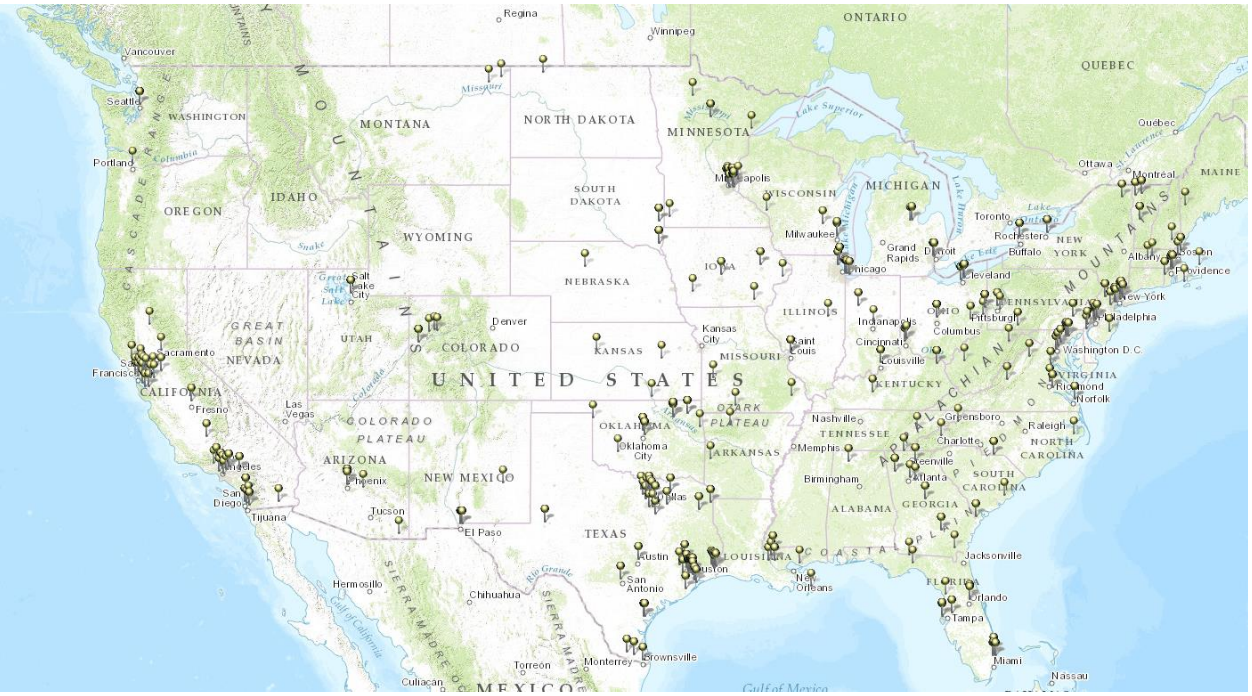
To estimate the ability of HEM-3 to replicate the 2011 observed ambient concentrations of air toxics, we conducted an annual operational model performance evaluation for air toxics used in the non-hybrid model calculation. Statistical assessments of modeled versus observed concentrations were paired in time and space and aggregated on an annual basis. Exhibit F-54 contains a list of air toxics evaluated in the non-hybrid model performance evaluation and the number of paired sites (based on completeness criteria of observations, Section F-20) used in the annual average. Exhibit F-55 is a map of the 2011 monitoring locations for non-hybrid air toxics. Paired values of the monitored and modeled data are provided in the spreadsheet "nonhybrid_polls-model_evaluation_paired.xlsx" in the SupplementalData folder.

Exhibit F-54. Non-hybrid Air Toxics Evaluated

Non-hybrid Air Toxic	Number of Paired Sites
4-Methyl-2-pentanone	35
Acetonitrile	36
Antimony (TSP) LC	11
Antimony PM_{10} LC	18
Bromomethane	26
Carbon disulfide	23
Carbon tetrachloride	163
Chloromethane	155
Cobalt (TSP) LC	12
Cobalt PM_{10} LC	15
Ethylbenzene	170

Non-hybrid Air Toxic	Number of Paired Sites
Hexane	125
Isopropylbenzene	23
Methyl tert-butyl ether	6
Propanal	68
Selenium (TSP) LC	11
Selenium PM_{10} LC	14
Selenium $PM_{2.5}$ LC	58
Styrene	71
Tribromomethane	5
Vinyl Acetate	25

Exhibit F-55. 2011 Monitoring Locations for the Evaluation of Non-hybrid Air Toxics



The exhibits below (Exhibits F-54 through F-59) are boxplots containing ratio comparisons of model-to-monitor (ambient observations) concentrations for the evaluated non-hybrid air toxics. All air toxics shown below (except hexane) showed model underpredictions compared to ambient measurements. The modeled and monitored data are provided in the SupplementalData folder in "nonhybrid_polls-model_evaluation_paired.xlsx"

Exhibit F-56. Metal HAPs (Antimony PM_{10} LC, Antimony TSP LC, Cobalt PM_{10} LC, Cobalt TSP LC, Selenium PM_{10} LC, Selenium $PM_{2.5}$ LC, and Selenium TSP LC): 2011 Model/Monitor Ratios

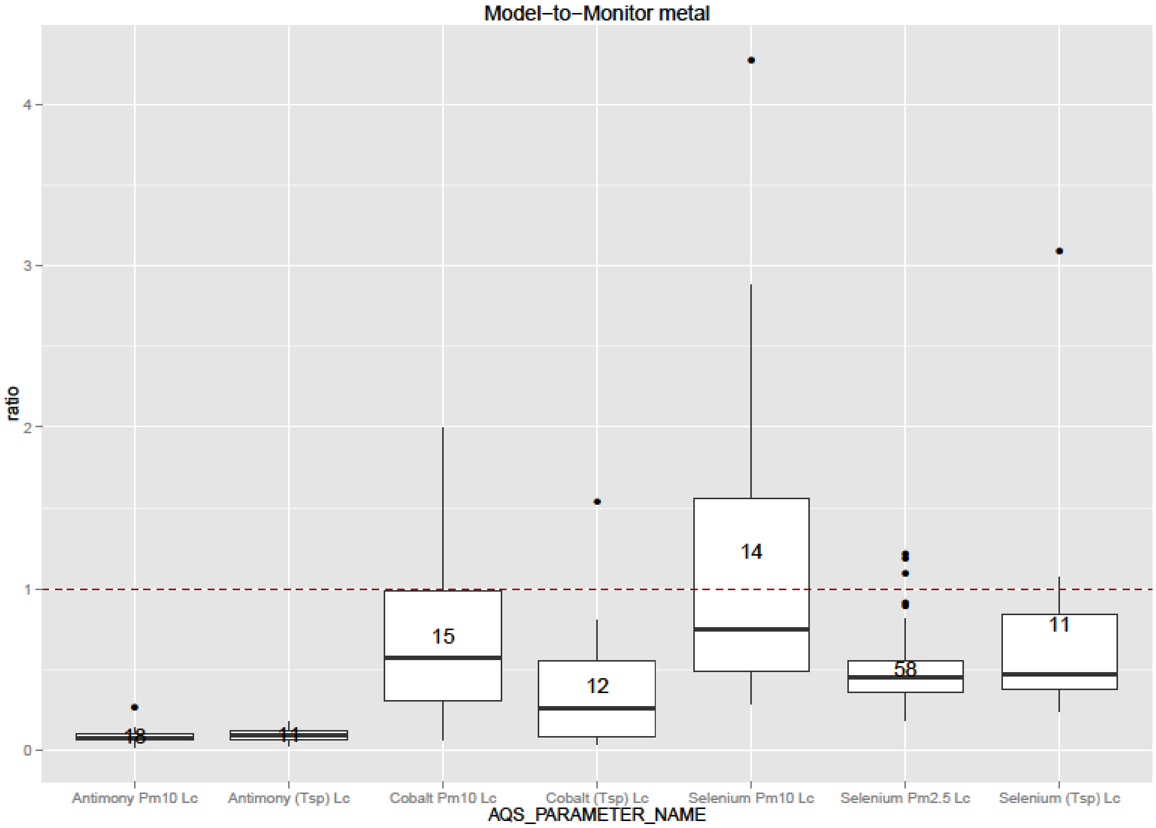


Exhibit F-57. Carbon Tetrachloride and Chloromethane: 2011 Model/Monitor Ratios

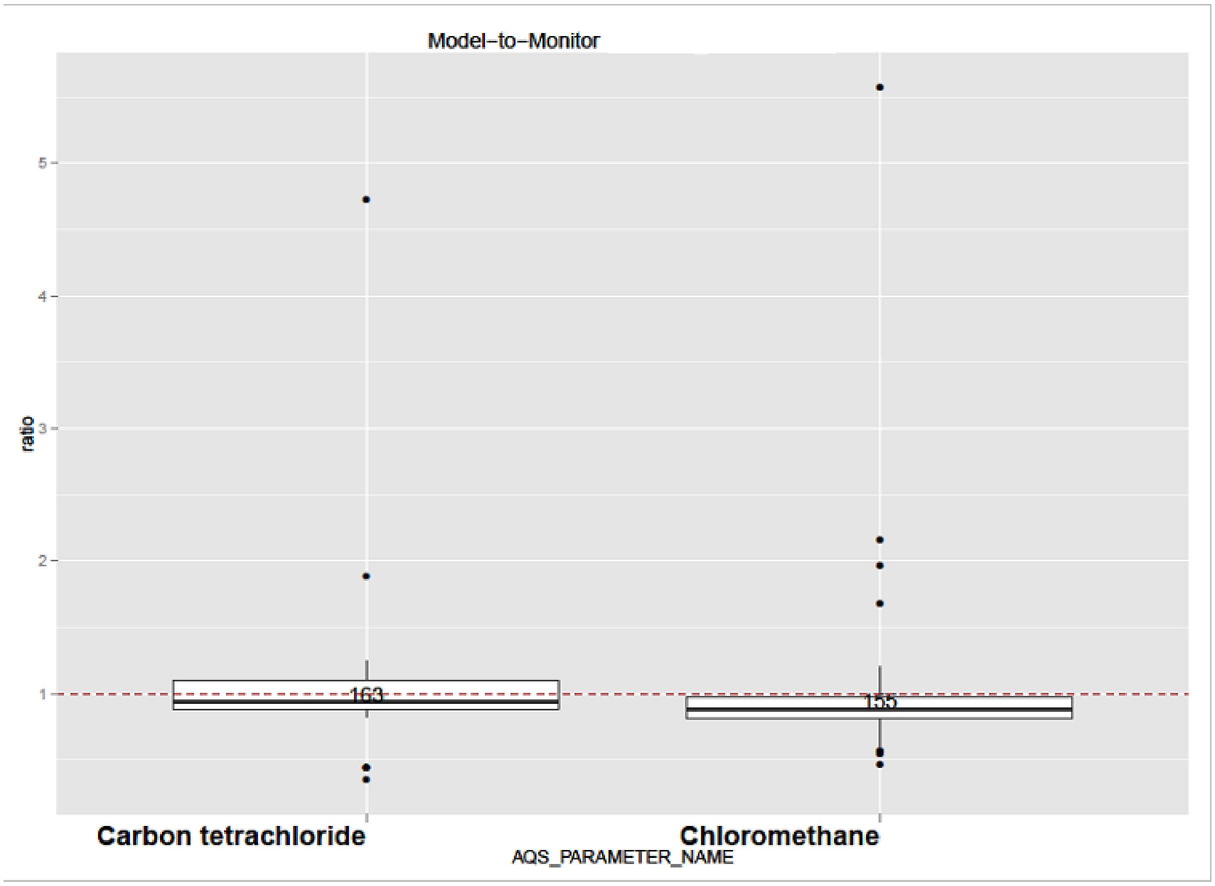


Exhibit F-58. Set #1 of Gaseous Air Toxics (Ethylbenzene, Hexane, Isopropylbenzene, Methyl Tert-butyl Ether, Propanal, and Styrene): 2011 Model/Monitor Ratios

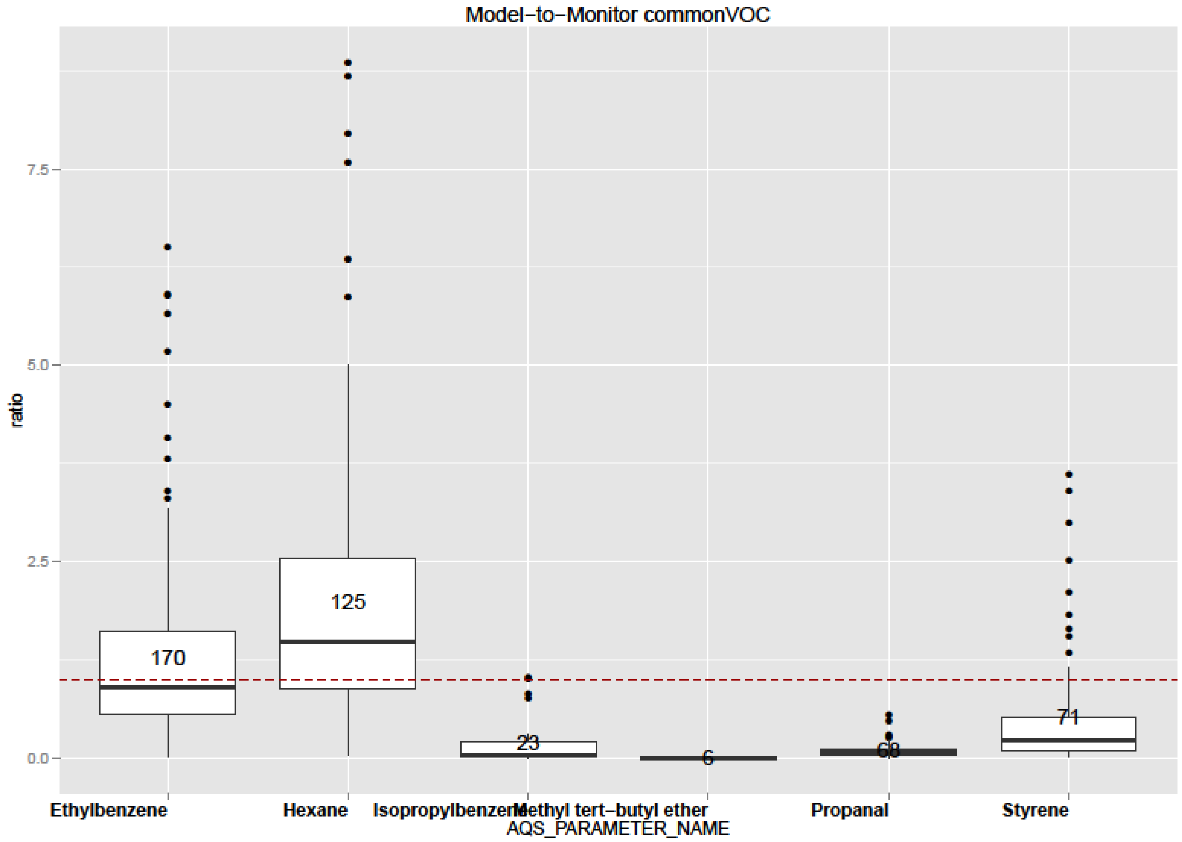
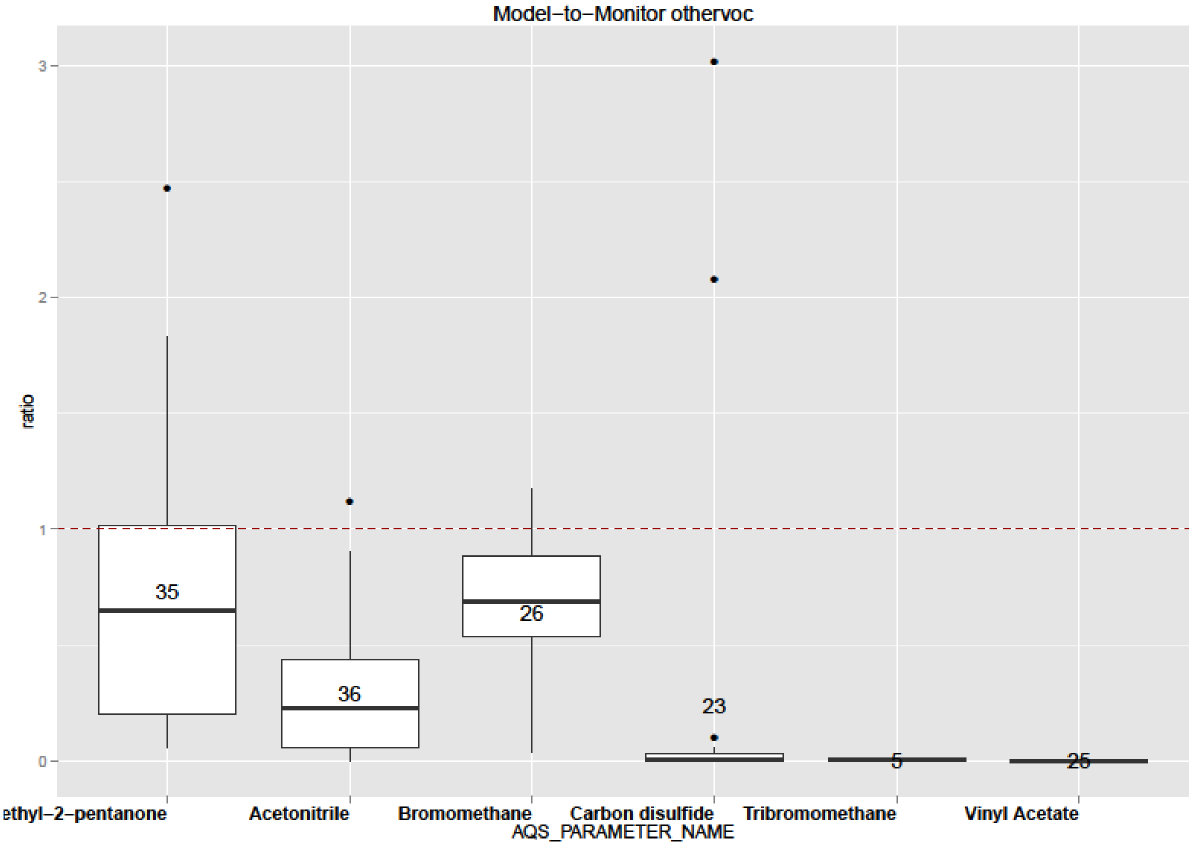


Exhibit F-59. Set #2 of Gaseous Air Toxics (4-methyl-2-pentanone, Acetonitrile, Bromomethane, Carbon Disulfide, Tribromomethane, and Vinyl Acetate): 2011 Model/Monitor Ratios



This page intentionally left blank.

Appendix G

Exposure Factors for the 2011 NATA

As noted in the memorandum contained within this Appendix (after Table G-1 below), ratios comparing exposure concentrations (ECs) predicted in HAPEM to air concentrations predicted in HEM-3 or CMAQ were developed for seven of the air toxic modeled for the 2011 NATA and applied to the remaining NATA air toxic s based on phase in ambient air (i.e., particulate, gaseous, or both; see Table 1 of the embedded memorandum in this appendix) and prevalence in emissions by NATA category. The ratios (or exposure factors) varied by air toxic and census tract, and the values were capped at maximum values determined as median plus one standard deviation (across all census tracts), with values 100 or larger treated as outliers and not part of the cap determinations (see Table 5 of the embedded memorandum in this appendix). Exhibit G-1 contains the factors (averaged across census tract) for each air toxic and each emission source sector (i.e., point, nonpoint, onroad mobile, nonroad mobile, and background), and the larger memorandum describing the HAPEM modeling for NATA follows Exhibit G-1.

Although these averaged factors were not actually used in NATA, they provide a general summary of the tract-level ratios that were used. These ratios account for the difference between the ambient outdoor concentration at a location and the EC that individuals were assumed to actually inhale in the risk assessment. Most of these averaged factors were less than 1, meaning that ECs tended to be lower than air concentrations. From among these averaged factors, predictions of EC ranged from 57 percent smaller than predictions of air concentration (for chromium (VI) from point sources) to 6 percent larger (for 1,3-butadiene from onroad mobile sources; note that exposure factors for 1,3-butadiene were not used as surrogates for any other NATA pollutants). The overall averaged exposure factor was 0.73 (i.e., ECs 27 percent smaller than air concentrations), likely due to the inability of many chemicals to penetrate efficiently into indoor environments.

A proximity term was required to adjust the predicted ambient air concentration (which was assumed to be representative of the census-tract centroid) to the level that we would expect immediately outside of the microenvironment. This proximity term was set to unity (i.e., 1) for most microenvironments. For the transportation-related microenvironments, however, ambient concentrations immediately outside the vehicle (i.e., very close to the pollutant source) were assumed to be considerably higher than at the predicted value at the census-tract centroid. Thus, we developed default proximity factors for HAPEM and used them for NATA. Separate factors also adjusted concentrations based on proximity to major roadways, as air concentrations will generally be higher near roadways than not near roadways.

Exhibit G-1. Averages of the Tract-level Exposure Factors Used in the 2011 NATA

Pollutant	Average Exposure Factor, by NATA Category [a,b]				Overall Average Across Categories at Left	Used as Surrogate For
	Point	Nonpoint	Onroad Mobile	Nonroad Mobile		
Benzene	0.88	0.87	1.04	0.88	0.92	All gaseous pollutants except 1,3-butadiene
1,3-butadiene	0.88	0.87	1.06	0.88	0.92	None
Coke oven emissions	1.00	--	--	--	1.00	None
Chromium (VI)	0.43	0.45	--	--	0.44	All particulate pollutants from point and nonpoint categories, except diesel PM
Diesel PM	--	0.67	0.56	0.46	0.56	None
Nickel	--	--	0.53	0.45	0.49	All particulate pollutants from mobile sources, except diesel PM
PAH, total	0.66	0.65	0.77	0.64	0.68	All pollutants that can be gaseous and particulate, except coke oven emissions

[a] Background and event sources not modeled in HAPEM

[b] Gray shading indicates a combination of pollutant and NATA category whose exposure factors were not applied as surrogates for other NATA pollutants, as noted in the "Used as Surrogate For" column.

MEMORANDUM

To: Ted Palma and Terri Hollingsworth
U.S. EPA, Office of Air Quality Planning and Standards

From: Chris Holder, Chris Stevens, Isaac Warren, Autumn Bordner, and Heidi Hubbard
ICF International

Date: 04/08/2015

Re: Running HAPEM7 for the 2011 National-scale Air Toxics Assessment (NATA)

The Hazardous Air Pollutant Exposure Model (HAPEM) is a U.S. Environmental Protection Agency (EPA) model used to perform screening-level assessments of long-term inhalation exposures to hazardous air pollutants (HAPs). The National-scale Air Toxics Assessment (NATA) is a nationwide EPA modeling assessment of the air concentrations, exposure concentrations, and potential human health cancer and non-cancer risks and hazards related to HAPs emitted by manmade and natural sources. For the purposes of the 2011 NATA, ICF ("we") used Version 7 of HAPEM (HAPEM7) with its default files and air-concentration outputs provided by EPA to model exposure concentrations for seven surrogate HAPs, stratified by census tract and source category. We used the modeled annual-exposure concentrations to estimate lifetime-exposure concentrations and divided those values by the corresponding annualized air concentrations (creating "exposure factors"). EPA can then use the exposure factors along with the NATA air concentrations to estimate exposure to each NATA HAP in each census tract.

In this memorandum, we discuss HAPEM7, how we identified the gas or particulate phase of the NATA HAPs, how EPA selected the seven HAPs to be modeled in HAPEM, how we set up the HAPEM runs (including linking to proximity and penetration factors), and how we developed the exposure factors. A tab-delimited text file ("NATA 2011 Exposure Factors_20150330") provided with this memorandum contains the exposure factors, and the Microsoft® Excel™ file "NATA 2011 Exposure Factors_Addl Information_20150330" contains additional information used to develop those factors.

1. Introduction to HAPEM and its Use in NATA

HAPEM is a model used by EPA to perform screening-level assessments of long-term inhalation exposures to HAPs. We completed updating HAPEM and its default, ancillary files in early 2015, as discussed in the ICF Memorandum “Updating the Hazardous Air Pollutant Exposure Model (HAPEM) for Use in the 2011 National-scale Air Toxics Assessment (NATA)” (from April 8, 2015, addressed to Ted Palma and Terri Hollingsworth of EPA’s Office of Air Quality Planning and Standards [OAQPS]). This latest version of HAPEM (HAPEM7)[1] estimates exposure concentrations using demographic and behavior data from the 2010 U.S. Census (covering all 50 states in the US, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands), compiled activity data from a recent version of EPA’s Consolidated Human Activity Database (CHAD), and other updated spatial data. The exposure concentrations are stratified by location (i.e., U.S. Census tract), time of day, age group, and the individual emission source categories and HAPs being modeled.

NATA is a nationwide modeling assessment of air concentrations, exposure concentrations, and potential, chronic human health risks and hazards associated with HAP emissions from man-made and naturally occurring sources. As described in this memorandum, we used HAPEM7 with its default files and with modeled air concentrations to model exposure concentrations for the 2011 NATA, for every census tract in the US, Puerto Rico, and the U.S. Virgin Islands.

2. Air Quality Inputs to HAPEM7

2.1. HAPs Assessed for the 2011 NATA

For the 2011 NATA, EPA compiled an inventory of the annual mass emitted of 273 HAPs from point, non-point, on-road mobile, and non-road mobile sources nationwide, typically at spatial resolutions of counties down to specific points. These HAPs are shown in Table 1, along with other information discussed later in this memorandum. Using appropriate hourly meteorology data and temporal emission profiles based on source category, EPA modeled these emissions in AERMOD (the atmospheric dispersion model developed by the American Meteorological Society and the EPA Regulatory Model Improvement Committee) and in CMAQ (EPA’s Community Multiscale Air Quality model). The air-modeling results were annual-average air concentrations at the level of census tracts, stratified by HAP, source type, and hour of day.

2.2. Specifying Chemical Phases for HAPEM Modeling

For all source types, HAPEM7 models microenvironment (ME) concentrations for three chemical phases. The model uses phase-specific penetration and proximity factors to estimate the ME chemical concentration for HAPs present in the gaseous (“G”) or particulate (“P”) phase as well as HAPs for which the phase can vary for typical atmospheric conditions (“G/P”). In addition, for on-road mobile sources of three HAPs (benzene, 1,3-butadiene, and diesel particulate matter [DPM]) and a fourth category

---

[1] As of January 20, 2015, HAPEM6 is available for download from http://www2.epa.gov/fera/download-hazardous-air-pollutant-exposure-model-hapem. We anticipate HAPEM7 and its User’s Guide will be made available by EPA online in Spring 2015.

representing all other HAPs, HAPEM7 uses chemical-specific penetration and proximity factors to estimate ME chemical concentrations.

Consequently, to prepare for NATA exposure modeling, we categorized each of the 273 NATA HAPs as G, P, or G/P based on available boiling-point data, as defined in Table 2. We provide each HAP's boiling point and assigned HAPEM7 HAP phase in Table 1. We obtained the vast majority of boiling-point values from either the Centers for Disease Control (CDC), the National Institutes of Health (NIH), or the Royal Society of Chemistry using their ChemSpider web site. These were judged to be the most reputable, comprehensive, and user-friendly sources of chemical boiling-point data readily available to us. Each allows the user to search by chemical name or Chemical Abstract Service number. For HAPs whose boiling points we could not identify using these three sources, we searched a variety of additional data sources. The source of each chemical's boiling-point value is provided in the "Source" column of Table 1.

We could not identify empirical boiling-point data for some of the HAPs. In many of these cases, we identified predictive boiling points from ChemSpider, which generates estimated boiling point using three software modules: EPA's EPIsuite, ACD/Labs Percepta Platform - PhysChem Module, and ChemAxon's predictive software platform. We typically selected the ACD/Labs values when available because these values were presented with confidence intervals and the conditions under which the boiling-point values were predicted (typically standard temperature and pressure); the other two platforms did not provide such information. If an ACD prediction was not available, we used the EPA EPIsuite value. Those boiling points that are predictive rather than empirical are flagged with a "P" in Table 2.

Note that the boiling point ranges in Table 1 have imprecise endpoints (e.g., the high end of boiling points for G HAPs covers a range of 240 to 260 °C). A relatively small number (i.e., 41) of NATA HAPs have boiling points within these imprecise endpoints, depending on the source of the data, meaning there was some uncertainty associated with assigning the phases for these HAPs. In order to make accurate designations, we conducted a literature review for each of these HAPs to identify relevant information regarding its typical physical state. For example, 1-nitropyrene has a boiling point of 445 °C, within the overlap of G/P and P boiling points. A review of the literature yielded several studies and reports identifying 1-nitropyrene as a particulate at typical atmospheric conditions, leading us to assign a designation of "P" to this HAP with a high degree of confidence. Where literature searches were uninformative, we assigned HAP phase based on the categorizations used for HAPEM5 to support the 1999 NATA.[2] The combination of the additional literature review and consultation of the HAPEM5 designations allowed us to make a reasonable phase designation for these HAPs. Nineteen HAPs have boiling points within the 240–260 °C range; based on the literature review and HAPEM5 designations, we categorized 15 as G and the remaining four as G/P. Twenty-two HAPs have boiling points within the 400–480 °C range; we categorized 16 as P and the remaining six as G/P.

In addition to the above 41 HAPs, boiling-point data were widely varying for three HAPs (see the "V" designations in the boiling-point-value column in Table 1). We categorized two of these HAPs, coke oven emissions and cyanide, as G/P. We categorized coke oven emissions based on an EPA characterization of

---

[2] The HAP categorizations for HAPEM5 for the purposes of the 1999 NATA were discussed in an ICF and TRJ Memorandum "Development of Penetration and Proximity Microenvironment Factor Distributions for the HAPEM5 in Support of the 1999 National-Scale Air Toxics Assessment (NATA)." from April 5, 2004, addressed to Ted Palma of EPA-OAQPS.

this pollutant as consisting of a mixture of particulates, volatiles, and semi-volatiles (http://www.epa.gov/ttnatw01/hlthef/cokeoven.html). We characterized cyanide based on the fact that cyanide is not typically found in isolation in nature, but rather in a variety of compounds, some of which are typically solid (e.g., calcium cyanide, sodium cyanide) and some of which are typically gaseous (e.g., hydrogen cyanide) (http://www.atsdr.cdc.gov/). The third HAP with widely-varying boiling-point data was diesel particulate matter and was assumed to be largely present as particulate.

For eight HAPs, boiling-point data were either unavailable or were ill-defined (see the “NA”, “D”, and “S” designations in the boiling-point-value column in Table 1); three were fibers, which we categorized as P, and we left the remaining five uncategorized pending potential resolution by EPA. The five HAPS currently without phase designations are as follows and are likely not emitted in large amounts:

- (Ethylenebis(Oxyethylenenitrilo)) Tetraacetic Acid
- Extractable Organic Matter (EOM)
- Propoxur
- Quinone
- Toxaphene

As noted previously, HAPEM7 does not use penetration and proximity factors specific to each NATA HAP, and the temporal emission profiles used in NATA air-concentration modeling vary only by the four source categories. Therefore, the level of effort to conduct exposure modeling on all 273 NATA HAPs is not justified. EPA identified a small subset of NATA HAPs for which to conduct HAPEM exposure modeling. NATA will use the HAPEM results of this modeled subset as-is, and these results will be used as surrogates for the remaining NATA HAPs not modeled in HAPEM. EPA used a subjective combination of decision points in identifying this subset, including

1. the subset must include at least one HAP per phase (i.e., at least one G HAP, one P HAP, and one G/P HAP);
2. collectively among the HAP(s) representing a phase, each emitting source category must be represented;
3. it is preferred that the selected HAPs pose high potential, relative risks to human health nationwide (using comparisons of air concentrations to health benchmarks, in a screening way without accounting for factors affecting exposure); and
4. it is preferred that the selected HAPs be emitted in many spatially-diverse locations across the US.

Using the above general criteria, EPA selected the seven HAPs listed below (and shown in Table 3) for exposure modeling.

- **Benzene** and **1,3-butadiene**, which are emitted by many processes (and all four modeled source categories) in nearly all U.S. locations. Benzene was selected to be the surrogate for all other G HAPs (EPA considers benzene modeling in NATA to be more reliable than 1,3-butadiene modeling).

- Unspeciated, generic PAHs ("**PAH, total**") are emitted by all four source categories and from a wide variety of processes, so it was selected to be the surrogate for all other G/P HAPs, except **coke oven emissions** which is a special case that was modeled by itself for point sources.
- **Chromium (VI)** is a highly toxic HAP that was selected as the surrogate for all other P HAPs emitted by point or non-point sources except for **DPM**, which was modeled as itself for non-point and mobile sources. Note that the NATA air-concentration modeling included chromium (VI) emissions from all four source categories, but its use as an exposure surrogate only included point and non-point sources because those are its major emitters.
- For P HAPs, besides DPM modeled as itself for non-point and mobile sources, and besides chromium (VI) being a surrogate for all other point and non-point P HAPs, **nickel** was selected as the surrogate for all other mobile-source P HAPs. Nickel is emitted by a variety of processes spread across the US. Note that NATA air-concentration modeling included nickel emissions from all four source categories, but its use as an exposure surrogate only included mobile sources because chromium (VI) was designated as the surrogate for point and non-point sources.

Whether a given HAP was modeled explicitly in HAPEM or is matched to a surrogate is indicated in Table 1 (final four columns). We used the air-concentration modeling outputs for these seven HAPs, stratified by source type, hour of day, and census tract, as the air-quality input files for seven HAPEM7 runs for NATA.

3. HAPEM7 Runs

3.1. Design

For each of the seven HAPEM7 runs (corresponding to each of the seven HAPs assessed for human exposure), we used the HAPEM7 default census- and CHAD-based files.[3] Each run assessed the 18 HAPEM7 MEs and all populated census tracts in the US, Puerto Rico, and the U.S. Virgin Islands. Though the air-quality inputs were by hour of day (i.e., 24 values per HAP, tract, and source category), HAPEM7 evaluated exposure in three-hour periods (i.e., 8 values per HAP, tract, and source category). We used the six default HAPEM7 age groups[4] and three day types.[5] We linked each HAP to its appropriate HAPEM7 penetration- and proximity-factors files, and we used the recommended setting of 30 replicates evaluated per HAP and tract.

3.2. Quality Control and Quality Assurance

We reviewed the HAPEM7 control files (i.e., "parameters" files) for accuracy, and then we reviewed the log, "counter," and "mistract" HAPEM7 output files to identify any potential errors in the modeling. We identified no errors in the inputs or outputs. We present below, and in Table 4, some statistics gleaned from the HAPEM7 log, counter, and mistract output files.

---

[3] The HAPEM7 default input files are described in the ICF Memorandum "Updating the Hazardous Air Pollutant Exposure Model (HAPEM) for Use in the 2011 National-scale Air Toxics Assessment (NATA)" from April 8, 2015, addressed to Ted Palma and Terri Hollingsworth of EPA-OAQPS.

[4] The HAPEM7 default age groups: 0–1, 2–4, 5–15, 16–17, 18–64, and 65 and older

[5] The HAPEM7 default day types: summer weekday, non-summer weekday, and weekend.

- 202 tracts (i.e., less than 1 percent of U.S. tracts) were not modeled for air concentrations because the population data EPA was using in air-concentration modeling indicated zero residents (the HAPEM7 population file had two to three residents for two of these tracts).
- An additional 377 tracts were not modeled in HAPEM7 because HAPEM7’s population data indicated zero residents.
- 1,027 areas modeled for air concentrations were not modeled in HAPEM7 because they were not census tracts (in addition to census tracts, EPA modeled air concentrations in some areas that corresponded to CMAQ grids; we did not model these areas for exposure).
- In total, 1,404 of the areas modeled for air concentrations were not modeled for exposure concentrations; 73,832 tracts were modeled for exposure concentrations for the 2011 NATA.

3.3. Post-processing into Exposure Factors

As has been done in previous NATAs, we utilized the HAPEM7 outputs for the seven assessed HAPs to estimate exposure factors that EPA will then apply to all HAPs assessed in NATA, based on HAP phase and source category. For each HAPEM7 run, and within that for each tract and source category, we calculated the estimated lifetime-average exposure concentration for each modeled replicate. We calculated these 70-year-average concentrations as the time-weighted average of exposures for the six HAPEM7 age groups, as shown below.

```
Lifetime average exposure conc. =
  [age group  0  –  1 exposure conc. × (  2 / 70 )]
+ [age group  2  –  4 exposure conc. × (  3 / 70 )]
+ [age group  5  – 15 exposure conc. × ( 11 / 70 )]
+ [age group 16  – 17 exposure conc. × (  2 / 70 )]
+ [age group 18  – 64 exposure conc. × ( 47 / 70 )]
+ [age group 65+      exposure conc. × (  5 / 70 )]
```

We then calculated the median lifetime-exposure concentration from the set of 30 replicates for each tract and source category for a given HAP. The sum of these medians across source categories yields the cumulative (i.e., from all modeled sources) “typical” lifetime-average exposure concentrations per HAP and tract.

For each assessed HAP, and then for each tract and source category (including the cumulative from all modeled sources), we divided these median lifetime-exposure concentrations by the corresponding annual air concentrations, resulting in an exposure factor. EPA can then multiply these exposure factors by the air concentrations of any appropriate HAP, resulting in estimated lifefime-exposure concentrations for that HAP. For example, for a given census tract, to estimate the exposure concentrations of a particular G/P HAP emitted by non-point sources, EPA will multiply the HAP’s non-point-source air concentrations by the non-point-source exposure factors for “PAH, total.”

The median and average exposure factors were between approximately 0.4 and 1.1 (larger factors typically for on-road mobile sources and gases; smaller factors typically for the other source categories and particulates). However, for a relatively small number of tracts, exposure factors were larger than 10 and, for approximately 100 tracts, exposure factors were larger than 100. Such large exposure factors are likely due to modeled people working in tracts with much larger air concentrations than their home tracts, so that their exposure factors account for home- and work-tract air quality but the air-concentration denominator in the exposure-factor calculation only accounts for home-tract air quality.

EPA considers these larger exposure factors to not reasonably represent average exposure scenarios across individual tracts. To ensure representative exposure factors, we have limited exposure factors to the maximum values shown in Table 5. These “caps”, specified per HAP and source category, correspond to the median exposure factor plus one standard deviation (taken across all tracts). The calculations for medians and standard deviations did not consider exposure factors 100 or larger, which we considered to be outliers. All applications of the exposure factors for the 2011 NATA use these caps, including for the HAPs explicitly modeled in HAPEM7.

A tab-delimited text file (“NATA 2011 Exposure Factors_20150330”) provided with this memorandum contains the exposure factors. The Microsoft® Excel™ file “NATA 2011 Exposure Factors_Addl Information_20150330” contains additional useful information such as the exposure and air concentrations as well as the calculations used to cap the exposure factors. The latter file includes warning flags for any situations where we did not calculate exposure factors. These situations correspond to air concentrations of 0, or exposure concentrations of 0 (indicating zero residents), or exposure concentrations not calculated (the 1,027 CMAQ grids). In all these flagged situations, we forced the exposure factor to be a value of 1, indicating that the air concentration equals the exposure concentration.

We quality-assured this post-processing by scrutinizing the SAS code used to accomplish it, spot-checking its calculations manually, and other broad checks to ensure all records were properly read in and the flagged records were properly processed.

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
147	(Ethylenebis(Oxyethylenenitrilo)) Tetraacetic Acid	✓				NA	NA	NA	NA			
255	1,1,2,2-Tetrachloroethane	✓	✓			147	CDC	G	Benz	Benz		
266	1,1,2-Trichloroethane	✓	✓			114	CDC	G	Benz	Benz		
119	1,1-Dimethyl Hydrazine	✓				64	CDC	G	Benz			
176	1,2,3,4,5,6-Hexachlorocyclohexane	✓	✓			323	CDC	G/P	PAH	PAH		
263	1,2,4-Trichlorobenzene	✓	✓			213	CDC	G	Benz	Benz		
92	1,2-Dibromo-3-Chloropropane	✓				196	CDC	G	Benz			
110	1,2-Dimethoxyethane	✓				82	NIH	G	Benz			
126	1,2-Diphenylhydrazine	✓				293	NIH	G/P	PAH			
128	1,2-Epoxybutane	✓	✓			63	NIH	G	Benz	Benz		
246	1,2-Propylenimine	✓				66	NIH	G	Benz			
41	1,3-Butadiene	✓	✓	✓	✓	138	NIH	G	✓	✓	✓	✓
98	1,3-Dichloropropene	✓	✓			108	NIH	G	Benz	Benz		
240	1,3-Propanesultone	✓				180	NIH	G	Benz			
94	1,4-Dichlorobenzene	✓	✓			173	CDC	G	Benz	Benz		
195	12-Methylbenz(a)Anthracene		✓			410 P	CS	P		Cr6		
203	1-Methylnaphthalene	✓				240	NIH	G	Benz			
205	1-Methylphenanthrene		✓			359	http://www.nature.nps.gov/hazardssafety/toxic/phen1met.pdf	G/P		PAH		
206	1-Methylpyrene					372	http://www.chemicalbook.com/ChemicalProductProperty_EN_CB7421679.htm	G/P				
219	1-Nitropyrene	✓				445 P	CS	P	Cr6			
166	2-(Hexyloxy)Ethanol	✓				258	NIH	G	Benz			
274	2,2,4-Trimethylpentane	✓	✓	✓	✓	99	NIH	G	Benz	Benz	Benz	Benz
264	2,4,4'-Trichlorobiphenyl (PCB-28)	✓	✓			164	CS	G	Benz	Benz		
268	2,4,5-Trichlorophenol	✓				247	NIH	G	Benz			
269	2,4,6-Trichlorophenol	✓	✓			246	NIH	G	Benz	Benz		

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
78	2,4-Dichlorophenoxy Acetic Acid	✓	✓			345 P	CS	G/P	PAH	PAH		
121	2,4-Dinitrophenol	✓	✓			312	CS	G/P	PAH	PAH		
124	2,4-Dinitrotoluene	✓	✓			300	NIH	G/P	PAH	PAH		
260	2,4-Toluene Diisocyanate	✓	✓			251	NIH	G	Benz	Benz		
7	2-Acetylaminofluorene	✓				400 P	CS	G/P	PAH			
42	2-Butoxyethyl Acetate	✓	✓			192	CDC	G	Benz	Benz		
58	2-Chloroacetophenone	✓	✓			244	CDC	G	Benz	Benz		
238	2-Chlorobiphenyl (PCB-1)	✓	✓			290	NIH	G/P	PAH	PAH		
63	2-Chloronaphthalene	✓	✓			256	http://www.chemicalbook.com/ChemicalProductProperty_EN_CB8854627.htm	G	Benz	Benz		
204	2-Methylnaphthalene	✓	✓	✓		241	http://www.speclab.com/compound/c91576.htm	G	Benz	Benz	Benz	
218	2-Nitropropane	✓	✓			121	CDC	G	Benz	Benz		
154	2-Propoxyethyl Acetate	✓				184	CS	G	Benz			
95	3,3'-Dichlorobenzidine	✓				400	NIH	G/P	PAH			
109	3,3'-Dimethoxybenzidine	✓				391	CS	G/P	PAH			
117	3,3'-Dimethylbenzidine	✓				300	CDC	G/P	PAH			
197	3-Methylcholanthrene	✓	✓			178	http://www.speclab.com/compound/c50328.htm	G	Benz	Benz		
96	4,4'-Dichlorobiphenyl (PCB-15)	✓	✓			144	CS	G	Benz	Benz		
199	4,4'-Methylenebis(2-Chloraniline)	✓				209	NIH	G	Benz			
202	4,4'-Methylenedianiline	✓	✓			397	CDC	G/P	PAH	PAH		
201	4,4'-Methylenediphenyl Diisocyanate	✓	✓			313	CS	G/P	PAH	PAH		
120	4,6-Dinitro-o-Cresol	✓				312	CDC	G/P	PAH			
13	4-Aminobiphenyl	✓				302	CDC	G/P	PAH			
114	4-Dimethylaminoazobenzene	✓				371	CS	G/P	PAH			
214	4-Nitrobiphenyl	✓				340	CDC	G/P	PAH			
217	4-Nitrophenol	✓	✓			279	NIH	G/P	PAH	PAH		

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
198	5-Methylchrysene	✓	✓			449 P	CS	P	Cr6	Cr6		
116	7,12-Dimethylbenz[a]Anthracene	✓	✓			122	http://www.sigmaaldrich.com/catalog/product/supelco/442425?lang=en&region=US	G	Benz	Benz		
90	7H-Dibenzo[c,g]carbazole	✓				544 P	CS	P	Cr6			
184	9-Methyl Anthracene		✓			196	CS	G				
1	Acenaphthene	✓	✓	✓	✓	279	NIH	G/P	PAH	PAH	PAH	PAH
2	Acenaphthylene	✓	✓	✓	✓	265	NIH	G/P	PAH	PAH	PAH	PAH
3	Acetaldehyde	✓	✓	✓	✓	20	CDC	G	Benz	Benz	Benz	Benz
4	Acetamide	✓	✓			165	CDC	G	Benz	Benz		
5	Acetonitrile	✓	✓			82	CDC	G	Benz	Benz		
6	Acetophenone	✓	✓			202	NIH	G	Benz	Benz		
8	Acrolein	✓	✓	✓	✓	53	CDC	G	Benz	Benz	Benz	Benz
9	Acrylamide	✓	✓			175	CDC	G	Benz	Benz		
10	Acrylic Acid	✓	✓			141	CDC	G	Benz	Benz		
11	Acrylonitrile	✓	✓			77	CDC	G	Benz	Benz		
12	Allyl Chloride	✓	✓			45	CDC	G	Benz	Benz		
14	Aniline	✓	✓			184	CDC	G	Benz	Benz		
16	Anthracene	✓	✓	✓	✓	342	NIH	G/P	PAH	PAH	PAH	PAH
17	Antimony	✓	✓			1,587	CDC	P	Cr6	Cr6		
18	Arsenic	✓	✓	✓	✓	612	CDC	P	Cr6	Cr6	Ni	Ni
19	Asbestos	✓	✓			600	CDC	P	Cr6	Cr6		
20	Benz[a]Anthracene	✓	✓	✓	✓	438	NIH	P	Cr6	Cr6	Ni	Ni
21	Benzene	✓	✓	✓	✓	80	CDC	G	✓	✓	✓	✓
22	Benzidine	✓				400	CDC	G/P	PAH			
23	Benzo(a)Fluoranthene	✓	✓			295	NIH	G/P	PAH	PAH		
24	Benzo(c)phenanthrene					430 P	CS	P				
26	Benzo(g,h,i)Fluoranthene	✓	✓			406 P	CS	P	Cr6	Cr6		
28	Benzo[a]Pyrene	✓	✓	✓	✓	360	http://www.speclab.com/compound/c50328.htm	G/P	PAH	PAH	PAH	PAH

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
29	Benzo[b]Fluoranthene	✓	✓	✓	✓	4,665 P	CS	P	Cr6	Cr6	Ni	Ni
25	Benzo[e]Pyrene	✓	✓			465 P	CS	P	Cr6	Cr6		
27	Benzo[g,h,i,]Perylene	✓	✓	✓	✓	550	NIH	P	Cr6	Cr6	Ni	Ni
30	Benzo[j]fluoranthene	✓				480 E	CS	P	Cr6			
31	Benzo[k]Fluoranthene	✓	✓	✓	✓	480	http://www.speclab.com/compound/c207089.htm	P	Cr6	Cr6	Ni	Ni
32	Benzofluoranthenes	✓	✓			406 P	CS	P	Cr6	Cr6		
33	Benzotrichloride	✓				221	NIH	G	Benz			
34	Benzyl Chloride	✓	✓			179	NIH	G	Benz	Benz		
35	Beryllium	✓	✓			2,500	NIH	P	Cr6	Cr6		
36	Beta-Propiolactone	✓				162	http://www.cdc.gov/niosh/docs/81-123/pdfs/0528.pdf	G	Benz			
37	Biphenyl	✓	✓			256	NIH	G	Benz	Benz		
38	Bis(2-Ethylhexyl)Phthalate	✓	✓			386	CDC	G/P	PAH	PAH		
39	Bis(Chloromethyl)Ether	✓				106	CDC	G	Benz			
40	Bromoform	✓	✓			149	CDC	G	Benz	Benz		
43	Butyl Carbitol Acetate	✓	✓			245	NIH	G	Benz	Benz		
44	Cadmium	✓	✓			765	CDC	P	Cr6	Cr6		
45	Calcium Cyanamide	✓	✓			>2,444	CDC	P	Cr6	Cr6		
46	Captan	✓	✓			314	CS	G/P	PAH	PAH		
47	Carbaryl	✓	✓			315	CS	G/P	PAH	PAH		
48	Carbazole	✓	✓			355	http://www.sigmaaldrich.com/catalog/product/sigma/c5132?lang=en&region=US	G/P	PAH	PAH		
49	Carbitol Acetate	✓				219	NIH	G	Benz			
50	Carbon Disulfide	✓	✓			47	CDC	G	Benz	Benz		
51	Carbon Tetrachloride	✓	✓			77	CDC	G	Benz	Benz		
52	Carbonyl Sulfide	✓	✓			-50	NIH	G	Benz	Benz		
53	Catechol	✓				245	CDC	G	Benz			
139	Cellosolve Acetate	✓	✓			145	CDC	G	Benz	Benz		
138	Cellosolve Solvent	✓	✓			124	CDC	G	Benz	Benz		

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
54	Chloramben	✓				312	CS	G/P	PAH			
55	Chlordane	✓				175	NIH	G	Benz			
56	Chlorine	✓	✓			-33	CDC	G	Benz	Benz		
57	Chloroacetic Acid	✓				106	CDC	G	Benz			
59	Chlorobenzene	✓	✓			132	CDC	G	Benz	Benz		
60	Chlorobenzilate	✓				146	NIH	G	Benz			
61	Chloroform	✓	✓			62	CDC	G	Benz	Benz		
62	Chloromethyl Methyl Ether	✓				59	CDC	G	Benz			
64	Chloroprene	✓	✓			59	CDC	G	Benz	Benz		
65	Chromic Acid (VI)	✓				250	CDC	G/P	PAH			
67	Chromium (VI)	✓	✓	✓	✓	2,642	CDC	P	✓	✓	Ni	Ni
66	Chromium III	✓	✓	✓	✓	2,672	http://books.google.com/books?id=SFD30BvPBhoC&pg=PA123&lpg=PA123&dq=chromium+III+melting+point&source=bl&ots=upHljDrKMy&sig=dISMKFL5z0sVI0z8Z4NhlsFHggE&hl=en&sa=X&ei=4nklVPLvJ4LS8AGbiYD4DA&ved=0CFkQ6AEwCQ#v=onepage&q=chromium%20III%20melting%20point&f=false	P	Cr6	Cr6	Ni	Ni
68	Chromium Trioxide	✓				250	CDC	G/P	PAH			
69	Chrysene	✓	✓	✓	✓	448	http://www.speclab.com/compound/c218019.htm	P	Cr6	Cr6	Ni	Ni
70	Coal Tar	✓				>250	http://www.inchem.org/documents/icsc/icsc/eics1415.htm	G/P	PAH			
71	Cobalt	✓	✓			3,100	CDC	P	Cr6	Cr6		
72	Coke Oven Emissions	✓				V	CDC	G/P	✓			
76	Cresol/Cresylic Acid (Mixed Isomers)	✓	✓			202	CDC	G	Benz	Benz		
77	Cumene	✓	✓			152	CDC	G	Benz	Benz		
283	Cyanide	✓	✓			V	CDC	G/P	PAH	PAH		
80	Decachlorobiphenyl (PCB-209)	✓	✓			460 P	CS	P	Cr6	Cr6		

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
81	Di(Ethylene Glycol Monobutyl Ether) Phthalate	✓				283	CDC	G/P	PAH			
82	Diazomethane	✓				-23	CDC	G	Benz			
83	Dibenz[a,h]acridine	✓				534 P	CS	P	Cr6			
85	Dibenzo[a,e]Pyrene	✓				552 P	CS	P	Cr6			
86	Dibenzo[a,h]Anthracene	✓	✓	✓	✓	262	http://www.sigmaaldrich.com/catalog/product/supelco/48574?lang=en&region=US	G/P	PAH	PAH	PAH	PAH
87	Dibenzo[a,h]Pyrene	✓				308 E	CS	G/P	PAH			
88	Dibenzo[a,i]Pyrene	✓				552 P	CS	P	Cr6			
84	Dibenzo[a,j]Acridine	✓				534 P	CS	P	Cr6			
89	Dibenzo[a,l]Pyrene	✓				552 P	CS	P	Cr6			
91	Dibenzofuran	✓	✓			287	NIH	G/P	PAH	PAH		
93	Dibutyl Phthalate	✓	✓			340	CDC	G/P	PAH	PAH		
97	Dichloroethyl Ether	✓				177	CDC	G	Benz			
99	Dichlorvos	✓				140 at 40 mmHG	NIH	G	Benz			
284	Diesel PM		✓	✓	✓	V	http://www.epa.gov/region1/eco/airtox/diesel.html	P		✓	✓	✓
100	Diethanolamine	✓	✓			268	NIH	G/P	PAH	PAH		
101	Diethyl Sulfate	✓				210	NIH	G	Benz			
103	Diethylene Glycol Diethyl Ether	✓				189 E	CS	G	Benz			
104	Diethylene Glycol Dimethyl Ether	✓				161 E	CS	G	Benz			
106	Diethylene Glycol Monobutyl Ether	✓	✓			230	NIH	G	Benz	Benz		
107	Diethylene Glycol Monoethyl Ether	✓	✓			196	NIH	G	Benz	Benz		
108	Diethylene Glycol Monomethyl Ether	✓	✓			194 E	CS	G	Benz	Benz		
112	Dimethyl Phthalate	✓	✓			284	NIH	G/P	PAH	PAH		

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
113	Dimethyl Sulfate	✓	✓			188	NIH	G	Benz	Benz		
118	Dimethylcarbamoyl Chloride	✓				167	NIH	G	Benz			
127	Epichlorohydrin	✓	✓			118	NIH	G	Benz	Benz		
130	Ethyl Acrylate	✓	✓			99	CDC	G	Benz	Benz		
131	Ethyl Benzene	✓	✓	✓	✓	136	CDC	G	Benz	Benz	Benz	Benz
132	Ethyl Carbamate	✓				185	NIH	G	Benz			
133	Ethyl Chloride	✓	✓			-139	NIH	G	Benz	Benz		
134	Ethylene Dibromide	✓	✓			131	CDC	G	Benz	Benz		
135	Ethylene Dichloride	✓	✓			83	CDC	G	Benz	Benz		
136	Ethylene Glycol	✓	✓			197	CDC	G	Benz	Benz		
140	Ethylene Glycol Methyl Ether	✓	✓			124	CDC	G	Benz	Benz		
141	Ethylene Glycol Monomethyl Ether Acetate	✓				145	CDC	G	Benz			
142	Ethylene Glycol Mono-Sec-Butyl Ether	✓				192	CS	G	Benz			
145	Ethylene Oxide	✓	✓			11	CDC	G	Benz	Benz		
146	Ethylene Thiourea	✓				230	CDC	G	Benz			
144	Ethyleneimine	✓				56	CDC	G	Benz			
148	Ethylidene Dichloride	✓	✓			-17	CDC	G	Benz	Benz		
149	Extractable Organic Matter (EOM)	✓				NA	NA	NA	NA			
150	Fine Mineral Fibers	✓				NA	http://www.usg.com/content/dam/USG_Marketing_Communications/united_states/product_promotional_materials/finished_assets/usg-mineral-wool-300a-msds-en-75850002.pdf	P	Cr6			
151	Fluoranthene	✓	✓	✓	✓	384	NIH	G/P	PAH	PAH	PAH	PAH
152	Fluorene	✓	✓	✓	✓	295	NIH	G/P	PAH	PAH	PAH	PAH
153	Formaldehyde	✓	✓	✓	✓	-21	CDC	G	Benz	Benz	Benz	Benz

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
154	Glycol Ethers	✓	✓			120-240	http://msdssearch.dow.com/PublishedLiteratureDOWCOM/dh_012d/0901b8038012d976.pdf?filepath=oxysolvents/pdfs/noreg/110-00977.pdf&fromPage=GetDoc	G	Benz	Benz		
155	Heptachlor	✓				392	CS	G/P	PAH			
156	Heptachlorobiphenyl	✓	✓			415 P	CS	P	Cr6	Cr6		
157	Hexachlorobenzene	✓	✓			325	NIH	G/P	PAH	PAH		
158	Hexachlorobiphenyl	✓	✓			396 P	CS	G/P	PAH	PAH		
159	Hexachlorobutadiene	✓	✓			215	CDC	G	Benz	Benz		
160	Hexachlorocyclopentadiene	✓	✓			238	CDC	G	Benz	Benz		
161	Hexachloroethane	✓				187	CDC	G	Benz			
162	Hexamethylene Diisocyanate	✓	✓			212	NIH	G	Benz	Benz		
163	Hexamethylphosphoramide	✓				233	NIH	G	Benz			
164	Hexane	✓	✓	✓	✓	69	CDC	G	Benz	Benz	Benz	Benz
167	Hydrazine	✓	✓			113	CDC	G	Benz	Benz		
168	Hydrochloric Acid	✓	✓			-85	CDC	G	Benz	Benz		
170	Hydrogen Cyanide	✓	✓			26	CDC	G	Benz	Benz		
169	Hydrogen Fluoride	✓	✓			19	CDC	G	Benz	Benz		
172	Hydroquinone	✓	✓			285	CDC	G/P	PAH	PAH		
173	Indeno[1,2,3-c,d]Pyrene	✓	✓	✓	✓	530	http://www.speclab.com/compound/c193395.htm	P	Cr6	Cr6	Ni	Ni
174	Isophorone	✓	✓			215	CDC	G	Benz	Benz		
175	Lead	✓	✓			1,740	CDC	P	Cr6	Cr6		
177	Maleic Anhydride	✓	✓			202	CDC	G	Benz	Benz		
178	Manganese	✓	✓	✓	✓	1,962	CDC	P	Cr6	Cr6	Ni	Ni
74	m-Cresol	✓	✓			202	CDC	G	Benz	Benz		
179	Mercury	✓	✓	✓	✓	356	CDC	G/P	PAH	PAH	PAH	PAH
180	Methanol	✓	✓			64	CDC	G	Benz	Benz		
182	Methoxychlor	✓				89	NIH	G	Benz			

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
183	Methoxytriglycol	✓				249	http://msdssearch.dow.com/PublishedLiteratureDOWCOM/dh_012d/0901b8038012d976.pdf?filepath=oxysolvents/pdfs/noreg/110-00977.pdf&fromPage=GetDoc	G	Benz			
185	Methyl Bromide	✓	✓			3	CDC	G	Benz	Benz		
187	Methyl Chloride	✓	✓			-98	CDC	G	Benz	Benz		
265	Methyl Chloroform	✓	✓			74	CDC	G	Benz	Benz		
189	Methyl Iodide	✓	✓			43	CDC	G	Benz	Benz		
190	Methyl Isobutyl Ketone	✓	✓			116	CDC	G	Benz	Benz		
191	Methyl Isocyanate	✓				39	CDC	G	Benz			
192	Methyl Methacrylate	✓	✓			101	CDC	G	Benz	Benz		
193	Methyl Tert-Butyl Ether	✓	✓			55	NIH	G	Benz	Benz		
194	Methylanthracene	✓				360 P	CS	G/P	PAH			
196	Methylbenzopyrene					479 P	CS	G/P				
198	Methylchrysene					449 P	CS	P				
200	Methylene Chloride	✓	✓			39	CDC	G	Benz	Benz		
188	Methylhydrazine	✓	✓			88	CDC	G	Benz	Benz		
279	m-Xylene	✓	✓	✓	✓	139	CDC	G	Benz	Benz	Benz	Benz
115	N,N-Dimethylaniline	✓	✓			192	CDC	G	Benz	Benz		
111	N,N-Dimethylformamide	✓	✓			153	CDC	G	Benz	Benz		
207	Naphthalene	✓	✓	✓	✓	260	CDC	G	Benz	Benz	Benz	Benz
165	N-Hexyl Carbitol	✓				260 E	CS	G	Benz			
208	Nickel	✓	✓	✓	✓	2,913	CDC	P	Cr6	Cr6	✓	✓
209	Nickel Oxide	✓				1,955	NIH	P	Cr6			
210	Nickel Refinery Dust	✓				2,730	http://www.cdc.gov/niosh/docs/81-123/pdfs/0445.pdf	P	Cr6			
213	Nitrobenzene	✓	✓			211	CDC	G	Benz	Benz		
222	N-Nitrosodimethylamine	✓				152	CDC	G	Benz			
223	N-Nitrosomorpholine	✓				224	NIH	G	Benz			
221	N-Nitroso-N-Methylurea	✓				164 P	CS	G	Benz			

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
15	o-Anisidine	✓				225	CDC	G	Benz			
73	o-Cresol	✓	✓			191	CDC	G	Benz	Benz		
261	o-Toluidine	✓	✓			200	CDC	G	Benz	Benz		
280	o-Xylene	✓	✓	✓	✓	144	CDC	G	Benz	Benz	Benz	Benz
224	PAH, total	✓	✓	✓	✓	240-400	http://www.epa.gov/reg3hwmd/bf-lr/regional/analytical/semi-volatile.htm	G/P	✓	✓	✓	✓
239	PAH/POM - Unspecified	✓	✓			100-450	http://www.epa.gov/reg3hwmd/bf-lr/regional/analytical/semi-volatile.htm	G/P	PAH	PAH		
225	Parathion	✓				375	CDC	G/P	PAH			
75	p-Cresol	✓	✓			202	CDC	G	Benz	Benz		
125	p-Dioxane	✓	✓			101	CDC	G	Benz	Benz		
226	Pentachlorobiphenyl	✓	✓			365 E	CS	G/P	PAH	PAH		
227	Pentachloronitrobenzene	✓	✓			328	NIH	G/P	PAH	PAH		
228	Pentachlorophenol	✓	✓			309	CDC	G/P	PAH	PAH		
229	Perylene	✓	✓			276	http://www.sigmaaldrich.com/catalog/product/aldrich/394475?lang=en&region=US	G/P	PAH	PAH		
230	Phenanthrene	✓	✓	✓	✓	340	NIH	G/P	PAH	PAH	PAH	PAH
231	Phenol	✓	✓			182	CDC	G	Benz	Benz		
232	Phenyl Cellosolve	✓	✓			245 E	CS	G	Benz	Benz		
234	Phosgene	✓	✓			8	CDC	G	Benz	Benz		
235	Phosphine	✓	✓			88	CDC	G	Benz	Benz		
236	Phosphorus	✓	✓			280	CDC	G/P	PAH	PAH		
237	Phthalic Anhydride	✓	✓			295	CDC	G/P	PAH	PAH		
238	Polychlorinated Biphenyls	✓	✓			365 E	CS	G/P	PAH	PAH		
233	p-Phenylenediamine	✓				267	CDC	G/P	PAH			
241	Propionaldehyde	✓	✓	✓	✓	48	NIH	G	Benz	Benz	Benz	Benz
242	Propoxur	✓				D	CDC	NA	NA			
243	Propyl Cellosolve	✓	✓			150	http://msdssearch.dow.com/PublishedLiteratureDOWCOM/dh_012d/0901b8038012d976.pdf?filepath=oxysolvents/pdfs/noreg/110-00977.pdf&fromPage=GetDoc	G	Benz	Benz		

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
244	Propylene Dichloride	✓	✓			97	CDC	G	Benz	Benz		
245	Propylene Oxide	✓	✓			34	CDC	G	Benz	Benz		
281	p-Xylene	✓	✓	✓	✓	138	CDC	G	Benz	Benz	Benz	Benz
247	Pyrene	✓	✓	✓	✓	404	NIH	P	Cr6	Cr6	Ni	Ni
248	Quinoline	✓	✓			238	NIH	G	Benz	Benz		
249	Quinone	✓				S	CDC	NA	NA			
150	Rockwool (Man-Made Fibers)	✓				NA	http://www.usg.com/content/dam/USG_Marketing_Communications/united_states/product_promotional_materials/finished_assets/usg-mineral-wool-300a-msds-en-75850002.pdf	P	Cr6			
250	Selenium	✓	✓			685	CDC	P	Cr6	Cr6		
150	Slagwool (Man-Made Fibers)	✓				NA	http://www.usg.com/content/dam/USG_Marketing_Communications/united_states/product_promotional_materials/finished_assets/usg-mineral-wool-300a-msds-en-75850002.pdf	P	Cr6			
251	Styrene	✓	✓	✓	✓	145	CDC	G	Benz	Benz	Benz	Benz
252	Styrene Oxide	✓	✓			194	http://www.sigmaaldrich.com/catalog/product/aldrich/s5006?lang=en&region=US	G	Benz	Benz		
254	Tetrachlorobiphenyl	✓	✓			360 P	CS	G/P	PAH	PAH		
256	Tetrachloroethylene	✓	✓			121	CDC	G	Benz	Benz		
257	Titanium Tetrachloride	✓	✓			136	http://www.sigmaaldrich.com/catalog/product/aldrich/697079?lang=en&region=US	G	Benz	Benz		
258	Toluene	✓	✓	✓	✓	111	CDC	G	Benz	Benz	Benz	Benz
259	Toluene-2,4-Diamine	✓				292	CDC	G/P	PAH			
262	Toxaphene	✓				D	CDC	NA	NA			
267	Trichloroethylene	✓	✓			87	CDC	G	Benz	Benz		
270	Triethylamine	✓	✓			89	CDC	G	Benz	Benz		
271	Triethylene glycol	✓	✓			285	NIH	G/P	PAH	PAH		
154	Triethylene Glycol Dimethyl Ether	✓				215 E	CS	G	Benz			
272	Trifluralin	✓	✓			140	http://www.speclab.com/compound/c1582098.htm	G	Benz	Benz		

Table 1. HAPs Assessed in the 2011 NATA, with their HAPEM7 HAP Phases

HAP Num. for NATA	HAP Name	Emission Source Modeled for NATA Air Concentrations[a]				Boiling Point		HAPEM7 HAP Phase[d]	Exposure Factors Developed for This HAP ("✓"), or Surrogate Used Instead (Surrogate HAP Name)[e]			
		P	NP	M-OR	M-NR	Value (C°)[b]	Source[c]		P	NP	M-OR	M-NR
273	Triglycol Monobutyl Ether	✓				278	NIH	G/P	PAH			
275	Vinyl Acetate	✓	✓			72	CDC	G	Benz	Benz		
276	Vinyl Bromide	✓				16	CDC	G	Benz			
277	Vinyl Chloride	✓	✓			-14	CDC	G	Benz	Benz		
278	Vinylidene Chloride	✓	✓			32	CDC	G	Benz	Benz		
282	Xylenes (Mixed Isomers)	✓	✓	✓	✓	139	NIH	G	Benz	Benz	Benz	Benz

[a] The emission sources modeled for air concentrations for each HAP in NATA are shown here for informational purposes. P =point; NP=non-point; M-OR=mobile on-road; M-NR=mobile non-road.

[b] D=decomposes; E=experimental; NA=not available; P=predicted; S=sublimes; V=varies depending on compound.

[c] CDC=http://www.cdc.gov/niosh/npg; CS=http://www.chemspider.com; NIH=http://pubchem.ncbi.nlm.nih.gov/.

[d] G=gaseous; G/P=gaseous or particulate depending on conditions; P=particulate; NA=unknown.

[e] Benz=benzene; PAH=PAH, total; Ni=Nickel; Cr6=Chromium (VI).

Table 2. Boiling-point Defintions Used to Classify HAPs for HAPEM7 Modeling for the 2011 NATA

HAPEM7 HAP Phase	Boiling-Point Range (°C)
G (Gaseous)	< 240-260
G/P (Either gaseous or particulate depending on conditions)	240-260 to 400-480
P (Particulate)	> 400-480

Source: Adapted from the “Classification of Inorganic Organic Pollutants” table at EPA’s Volatile Organic Compound page (available as of February 12, 2015 at http://www.epa.gov/iaq/voc2.html), as adapted from: World Health Organization, 1989. "Indoor air quality: organic pollutants." Report on a WHO Meeting, Berlin, 23-27 August 1987. EURO Reports and Studies 111. Copenhagen, World Health Organization Regional Office for Europe.

Table 3. HAPs Modeled in HAPEM7 for the 2011 NATA

NATA HAP	HAPEM7 HAP Phase[b]	Emission Source Modeled for NATA Exposure Concentrations[a]			
		P	NP	M-OR	M-NR
Benzene	G	✓	✓	✓	✓
1,3-butadiene	G	✓	✓	✓	✓
Coke oven emissions	G/P	✓			
PAH, total (i.e., aggregate mass of polycyclic aromatic hydrocarbons, where congeners were not specified)	G/P	✓	✓	✓	✓
Chromium (VI) (i.e., compounds of hexavalent chromium)	P	✓	✓		
Diesel PM (i.e., Diesel particulate matter, or DPM)	P		✓	✓	✓
Nickel	P			✓	✓

[a] For PAH total, chromium (VI), and nickel, we did not model exposure concentrations for the full set of source categories they were modeled with for air concentrations. As discussed in the text, EPA selected these seven HAPs to model and these specific source categories per HAP. P=point; NP=non-point; M-OR=mobile on-road; M-NR=mobile non-road.

[b] G=gaseous; G/P=gaseous or particulate depending on conditions; P=particulate.

Table 4. Modeling Statistics for the 2011 NATA HAPs Modeled in HAPEM7

<table>
<tr><th></th><th></th><th></th><th>Benzene</th><th>1,3-butadiene</th><th>Coke oven emissions</th><th>Chromium (VI)</th><th>Diesel PM</th><th>Nickel</th><th>PAH, total</th></tr>
<tr><td rowspan="3">In HAPEM7:</td><td colspan="2">Num. Counties</td><td colspan="7">3,224</td></tr>
<tr><td colspan="2">Num. Tracts</td><td colspan="7">74,034</td></tr>
<tr><td colspan="2">Num. Tracts with 0 Population (i.e., not modeled)[a]</td><td colspan="7">579</td></tr>
<tr><td rowspan="6">In 2011 NATA:</td><td rowspan="4">Num. Modeled:</td><td>Counties</td><td colspan="7">3,224</td></tr>
<tr><td>"Tracts" For Air Concentrations[b]</td><td colspan="7">74,859</td></tr>
<tr><td>"Tracts" For Air Concentrations but not Exposure[b]</td><td colspan="7">1,027</td></tr>
<tr><td>Tracts For Exposure</td><td colspan="7">73,832</td></tr>
<tr><td rowspan="2">Num. Not Modeled:</td><td>Tracts[a]</td><td colspan="7">202 for air concentrations, an additional 377 for exposure</td></tr>
<tr><td>Instances of a Work Tract[c]</td><td>3,202</td><td>3,202</td><td>3,149</td><td>3,149</td><td>3,149</td><td>3,149</td><td>3,149</td></tr>
</table>

[a] EPA did not model air concentrations for 202 tracts which had zero residents according to the population data EPA was using at that time. In the HAPEM7 population data, two of these tracts have two to three residents, but they were not modeled in HAPEM7 because they were not modeled in AERMOD and CMAQ; the other 200 tracts had zero residents in HAPEM7. The HAPEM7 population data indicate that another 379 tracts also have zero residents; though EPA modeled air concentrations for these tracts, HAPEM7 did not model these tracts.

[b] Air-concentration modeling included census tracts and some areas unrelated to census tracts. These non-tract areas were not modeled for exposure.

[c] Each home tract's collection of work tracts were randomly sampled with each run of HAPEM. A tract can be a work tract for multiple home tracts, and thus it can be sampled multiple times as a work tract. Work tracts were not modeled for exposure if they had no air concentrations; tracts had no air concentrations if they had no people in them according to the 2010 Census.

Table 5. “Caps” Applied to Exposure Factors for the 2011 NATA

HAP	Point	Non-point	Mobile On-road	Mobile Non-road	Total
Benzene	1.00	0.98	1.28	0.97	0.98
1,3-butadiene	1.09	0.97	1.30	0.99	1.00
Coke oven emissions	1.02	NA	NA	NA	1.02
PAH, total	0.93	0.71	0.89	0.73	0.73
Chromium (VI)	0.53	0.55	NA	NA	0.52
Diesel PM	NA	1.94	0.66	0.57	0.57
Nickel	NA	NA	0.64	0.54	0.58

This page intentionally left blank.

Appendix H

Toxicity Values Used in the 2011 NATA

Exhibit H-1 contains the toxicity values and supporting information for cancer and noncancer effects used in the 2011 NATA. The “target organ” column contains the organs or organ systems adversely affected at the lowest dose in human or animal studies of noncancer effects. Hazard indices were calculated only for the respiratory system (see Sections 5.2.2 and 6.3 of this document for the definitions of hazard quotients and hazard indices and an explanation of how they are used in NATA). Other information on individual substances is shown in footnotes. Abbreviations used for the sources of the unit risk estimates (UREs) and reference concentrations (RfCs) are as follows:

- IRIS = Integrated Risk Information System
- ATSDR = Agency for Toxic Substances and Disease Registry
- CAL = California Office of Environmental Health Hazard Assessment
- HEAST = EPA Health Effects Assessment Tables
- OAQPS = EPA Office of Air Quality Planning and Standards

Exhibit H-1. Toxicity Values Used in the 2011 NATA

NATA Pollutant	CAS Number	Inhalation Unit Risk Estimate (URE), 1/(µg/m^3)		Reference Concentration (RfC), mg/m^3		Target Organ(s) or System(s) Noncancer Impacts [a]
		Value	Source	Value	Source	
1,1,1-Trichloroethane	71556			5	IRIS	neurological
1,1,2,2-Tetrachloroethane	79345					
1,1,2-Trichloroethane	79005	0.000016	IRIS	0.4	CAL	liver
1,1-Dimethylhydrazine	57147					
1,2,3,4,5,6-Hexachlorocyclyhexane [b]	58899	0.00053	IRIS	0.002	CAL	liver reproductive kidney
1,2,4-Trichlorobenzene	120821			0.2	HEAST	liver
1,2-Dibromo-3-Chloropropane	96128	0.002	CAL	0.0002	IRIS	reproductive
1,2-Diphenylhydrazine	122667	0.00022	IRIS			
1,2-Epoxybutane	106887			0.02	IRIS	respiratory
1,2-Propyleneimine	75558					
1,3-Butadiene	106990	0.00003	IRIS	0.002	IRIS	reproductive
1,3-Dichloropropene	542756	0.000004	IRIS	0.02	IRIS	respiratory
1,3-Propane Sultone	1120714	0.00069	CAL			
1,4-Dichlorobenzene	106467	0.000011	CAL	0.8	IRIS	liver
1,4-Dioxane	123911	0.000005	IRIS	0.03	IRIS	respiratory liver
2,2,4-Trimethylpentane	540841					
2,4,5-Trichlorophenol	95954					
2,4,6-Trichlorophenol	88062	0.0000031	IRIS			
2,4-D, salts and esters	94757					
2,4-Dinitrophenol	51285					
2,4-Dinitrotoluene	121142	0.000089	CAL	0.007	CAL	liver neurological
2,4-Toluene Diamine	95807	0.0011	CAL			
2,4-Toluene Diisocyanate	584849	0.000011	CAL	0.00007	IRIS	respiratory
2-Acetylaminofluorene	53963	0.00208	CAL			
2-Chloroacetophenone	532274			0.00003	IRIS	respiratory
2-Nitropropane	79469	0.0000056 [c]	EPA OAQPS	0.02	IRIS	liver
3,3'-Dichlorobenzidine	91941	0.00034	CAL			
3,3'-Dimethoxybenzidine	119904					
3,3'-Dimethylybenzidine	119937					
4,4'-Methylene Bis(2-Chloroaniline)	101144	0.00043	CAL			
4,4'-Methylenedianiline	101779	0.00046	CAL	0.02	CAL	ocular
4,4'-Methylenediphenyl Diisocyanate (MDI)	101688			0.0006	IRIS	respiratory
4,6-Dinitro-o-Cresol (Including Salts)	534521					
4-Aminobiphenyl	92671					
4-Dimethylaminoazobenzene	60117	0.0013	CAL			
4-Nitrobiphenyl	92933					
4-Nitrophenol	100027					
Acetaldehyde	75070	0.0000022	IRIS	0.009	IRIS	respiratory
Acetamide	60355	0.00002	CAL			
Acetonitrile	75058			0.06	IRIS	whole body
Acetophenone	98862					

NATA Pollutant	CAS Number	Inhalation Unit Risk Estimate (URE), $1/(\mu g/m^3)$		Reference Concentration (RfC), mg/m^3		Target Organ(s) or System(s) Noncancer Impacts [a]
		Value	Source	Value	Source	
Acrolein	107028			0.00002	IRIS	respiratory
Acrylamide	79061	0.00016 [d]	IRIS	0.006	IRIS	neurological
Acrylic Acid	79107			0.001	IRIS	respiratory
Acrylonitrile	107131	0.000068	IRIS	0.002	IRIS	respiratory
Allyl Chloride	107051	0.000006	CAL	0.001	IRIS	neurological
Aniline	62533	0.0000016	CAL	0.001	IRIS	spleen
Anisidine	90040					
Antimony Compounds	7440360			0.0002	IRIS	respiratory
Arsenic Compounds(inorganic including Arsine)	7440382	0.0043	IRIS	0.000015	CAL	developmental
Benzene	71432	0.0000078 [e]	IRIS	0.03	IRIS	immune
Benzidine	92875	0.1072 [d]	IRIS	0.01	CAL	liver neurological
Benzotrichloride	98077					
Benzyl Chloride	100447	0.000049	CAL			
Beryllium Compounds	7440417	0.0024	IRIS	0.00002	IRIS	respiratory
Beta-Propiolactone	57578					
Biphenyl	92524					
Bis(2-Ethylhexyl)Phthalate (DEHP)	117817	0.0000024	CAL	0.01	CAL	respiratory liver
Bis(Chloromethyl) Ether	542881	0.062	IRIS			
Bromoform	75252	0.0000011	IRIS			
Cadmium Compounds	7440439	0.0018	IRIS	0.00001	ATSDR	kidney
Calcium Cyanamide	156627			0.0008	IRIS	neurological thyroid
Captan	133062					
Carbaryl	63252					
Carbon Disulfide	75150			0.7	IRIS	neurological
Carbon Tetrachloride	56235	0.000006		0.1	IRIS	liver
Carbonyl Sulfide	463581			0.163	EPA ORD [f]	neurological
Catechol	120809					
Chloramben	133904					
Chlordane	57749	0.0001	IRIS	0.0007	IRIS	liver
Chlorine	7782505			0.00015	ATSDR	respiratory
Chloroacetic Acid	79118					
Chlorobenzene	108907			1	CAL	liver reproductive kidney
Chlorobenzilate	510156	0.000078	HEAST			
Chloroform	67663			0.098	ATSDR	liver
Chloromethyl Methyl Ether	107302					
Chloroprene	126998	0.00048 [d]	IRIS	0.02	IRIS	respiratory
Chromium VI (Hexavalent) [g]	Multiple	0.012	IRIS	0.0001	IRIS	respiratory
Cobalt Compounds	7440484			0.0001	ATSDR	respiratory
Coke Oven Emissions	NA	0.00099 [d]	IRIS			
Cresol/Cresylic Acid (Mixed Isomers) [h]	Multiple			0.6	CAL	neurological whole body
Cumene	98828			0.4	IRIS	kidney endocrine

NATA Pollutant	CAS Number	Inhalation Unit Risk Estimate (URE), $1/(\mu g/m^3)$		Reference Concentration (RfC), mg/m^3		Target Organ(s) or System(s) Noncancer Impacts [a]
		Value	Source	Value	Source	
Cyanide Compounds	57125			0.0008	IRIS	neurological thyroid
Dibenzofuran	132649					
Dibutylphthalate	84742					
Dichloroethyl Ether (Bis[2-Chloroethyl]Ether)	111444	0.00033	IRIS			
Dichlorvos	62737			0.0005	IRIS	neurological
Diesel Particulate Matter				0.005	IRIS	respiratory
Diethanolamine	111422			0.003	CAL	respiratory
Diethyl Sulfate	64675					
Dimethyl Formamide	68122			0.03	IRIS	liver
Dimethyl Phthalate	131113					
Dimethyl Sulfate	77781					
Dimethylcarbamoyl Chloride	79447					
Epichlorohydrin	106898	0.0000012		0.001	IRIS	respiratory
Ethyl Acrylate	140885					
Ethyl Carbamate (Urethane) Chloride (Chloroethane)	51796	0.000464 [d]				
Ethyl Chloride	75003			10	IRIS	developmental
Ethylbenzene	100414	0.0000025	CAL	1	IRIS	developmental
Ethylene Dibromide (Dibromoethane)	106934	0.0006	IRIS	0.009	IRIS	respiratory reproductive
Ethylene Dichloride (1,2-Dichloroethane)	107062	0.000026	IRIS	2.4	ATSDR	liver
Ethylene Glycol	107211			0.4	CAL	respiratory
Ethylene Oxide	75218	0.000088	CAL	0.03	CAL	neurological
Ethylene Thiourea	96457	0.000013	CAL	0.003	CAL	endocrine
Ethyleneimine (Aziridine)	151564					
Ethylidene Dichloride (1,1-Dichloroethane)	75343	0.0000016	CAL	0.5	HEAST	kidney
Formaldehyde	50000	0.000013	IRIS	0.0098	ATSDR	respiratory
Glycol Ethers [i]	Multiple			0.02	IRIS	reproductive
Heptachlor	76448	0.0013	IRIS			
Hexachlorobenzene	118741	0.00046	IRIS	0.003	CAL	liver
Hexachlorobutadiene	87683	0.000022	IRIS	0.09	CAL	reproductive
Hexachlorocyclopentadiene	77474			0.0002	IRIS	respiratory
Hexachloroethane	67721			0.03	IRIS	liver neurological kidney
Hexamethylene Diisocyanate	822060			0.00001	IRIS	respiratory
Hexane	110543			0.7	IRIS	neurological
Hydrazine	302012	0.0049	IRIS	0.0002	CAL	liver thyroid
Hydrochloric Acid (Hydrogen Chloride [Gas Only])	7647010			0.02	IRIS	respiratory
Hydrogen Fluoride (Hydrofluoric Acid)	7664393			0.014	CAL	skeletal
Hydroquinone	123319					
Isophorone	78591			2	CAL	liver developmental

NATA Pollutant	CAS Number	Inhalation Unit Risk Estimate (URE), 1/(µg/m^3)		Reference Concentration (RfC), mg/m^3		Target Organ(s) or System(s) Noncancer Impacts [a]
		Value	Source	Value	Source	
Lead Compounds [j]	7439921			0.00015	EPA OAQPS	neurological developmental
Maleic Anhydride	108316			0.0007	CAL	respiratory
Manganese Compounds	7439965			0.0003	ATSDR	neurological
Mercury Compounds	7439976			0.0003 [k]	IRIS	neurological
Methanol	67561			20	IRIS	developmental
Methoxychlor	72435					
Methyl Bromide (Bromomethane)	74839			0.005	IRIS	respiratory
Methyl Chloride (Chloromethane)	74873			0.09	IRIS	neurological
Methyl Iodide (Iodomethane)	74884					
Methyl Isobutyl Ketone (Hexone)	108101			3	IRIS	developmental
Methyl Isocyanate	624839			0.001	CAL	respiratory whole body
Methyl Methacrylate	80626			0.7	IRIS	respiratory
Methyl Tert-Butyl Ether	1634044	0.00000026	CAL	3	IRIS	liver kidney ocular
Methylene Chloride	75092	0.000000016 [d]	IRIS	0.6	IRIS	respiratory liver
Methylhydrazine	60344					
N,N-Dimethylaniline	121697					
Naphthalene	91203	0.000034	CAL	0.003	IRIS	respiratory
Nickel Compounds [l]	1313991	0.00048	EPA OAQPS	0.00009	ATSDR	respiratory immune
Nitrobenzene	98953	0.00004	IRIS	0.009	IRIS	respiratory
N-Nitrosodimethylamine	62759	0.022 [d]	IRIS			
N-Nitrosomorpholine	59892	0.0019	CAL			
N-Nitroso-N-Methylurea	684935					
o-Toluidine	95534	0.000051	CAL			
PAH_000E0 (PAHPOM) [m]	Multiple					
PAH_176E5 (PAHPOM)	Multiple	0.0000176 [d]	EPA OAQPS			
PAH_880E5 (PAHPOM)	Multiple	0.000088 [d]	EPA OAQPS			
PAH_176E4 (PAHPOM)	Multiple	0.000176 [d]	EPA OAQPS			
PAH_176E3 (PAHPOM)	Multiple	0.00176 [d]	EPA OAQPS			
PAH_192E3 (PAHPOM)	Multiple	0.00192 [d]	EPA OAQPS			
PAH_101E2 (PAHPOM)	Multiple	0.01008 [d]	EPA OAQPS			
PAH_176E2 (PAHPOM)	Multiple	0.0176 [d]	EPA OAQPS			
PAH_114E1 (PAHPOM)	Multiple	0.1136 [d]	EPA OAQPS			
Parathion	56382					

NATA Pollutant	CAS Number	Inhalation Unit Risk Estimate (URE), 1/(µg/m^3)		Reference Concentration (RfC), mg/m^3		Target Organ(s) or System(s) Noncancer Impacts [a]
		Value	Source	Value	Source	
Pentachloronitrobenzene (Quintobenzene)	82688					
Pentachlorophenol	87865	0.0000051	CAL	0.1	CAL	liver kidney
Phenol	108952			0.2	CAL	liver
Phosgene	75445			0.0003	IRIS	respiratory
Phosphine	7803512			0.0003	IRIS	whole bod
Phosphorus	7723140					
Phthalic Anhydride	85449			0.02	CAL	respiratory ocular
Polychlorinated Biphenyls (Aroclors)	1336363	0.0001	IRIS			
p-Phenylenediamine	106503					
Propionaldehyde	123386			0.008	IRIS	respiratory
Propoxur (Baygon)	114261					
Propylene Dichloride (1,2-Dichloropropane)	78875			0.004	IRIS	respiratory
Propylene Oxide	75569	0.0000037	IRIS	0.03	IRIS	respiratory
Quinoline	91225					
Quinone (P-Benzoquinone)	106514					
Selenium Compounds	7782492			0.02	CAL	liver neurological hematologic
Styrene	100425			1	IRIS	neurological
Styrene Oxide	96093			0.006	CAL	respiratory
Tetrachloroethylene	127184	0.00000026	IRIS	0.04	IRIS	neurological
Titanium Tetrachloride	7550450			0.0001	ATSDR	respiratory
Toluene	108883			5	IRIS	neurological
Toxaphene (Chlorinated Camphene)	8001352	0.00032	IRIS			
Trichloroethylene	79016	0.0000048 [n]	IRIS	0.002	IRIS	liver neurological developmental reproductive kidney immune
Triethylamine	121448			0.007	IRIS	respiratory
Trifluralin	1582098					
Vinyl Acetate	108054			0.2	IRIS	respiratory
Vinyl Bromide	593602	0.000032	HEAST	0.003	IRIS	liver
Vinyl Chloride	75014	0.0000088	IRIS	0.1	IRIS	liver
Vinylidene Chloride	75354			0.2	IRIS	liver
Xylenes [o]	Multiple			0.1	IRIS	neurological

[a] For pollutants with more than one target organ or system listed, the order presented in this table does not represent priority or significance of the noncancer impact.

[b] Includes all 4 lindane isomers. The modeling used the toxicity values of the most toxic isomer for cancer (CAS 319857) to estimate risk.

[c] The URE for 2-nitropropane derived by the Health Council of the Netherlands in 1999 was used in preference to the value in the Health Effects Assessment Summary Tables, which does not reflect the most recent studies and analysis methods.

[d] This carcinogen acts via a mutagenic mode of action; therefore, the URE was adjusted by factor of 1.6 to account for the increased risk during childhood exposures.

NATA Pollutant	CAS Number	Inhalation Unit Risk Estimate (URE), 1/(μg/m^3)		Reference Concentration (RfC), mg/m^3		Target Organ(s) or System(s) Noncancer Impacts [a]
		Value	Source	Value	Source	

[e] The IRIS assessment for benzene contains a range of UREs for inhalation exposure. The values that bracket this range are based on different interpretations of the human-exposure information. As a health-protective national screening assessment, NATA used the upper end of the range.

[f] A chronic screening level of 0.163 mg/m^3 was developed for carbonyl sulfide by EPA ORD from a No Observed Adverse Effects Level of 200 ppm based on brain lesions and neurophysiological alteration in rodents.

[g] All hexavalent chromium compounds (including chromium (IV) trioxide) were modeled using the toxicity values for hexavalent chromium.

[h] The individual cresol isomers were combined and noncancer impacts were estimated using the RfC for their mixture.

[i] The RfC for ethylene glycol methyl ether was used for all glycol eithers.

[j] The RfC for lead compounds is equivalent to the lead NAAQS.

[k] The RfC for mercury compounds is the value derived by IRIS for elemental mercury.

[l] The IRIS assessments for nickel compounds provided a range of plausible UREs. NATA used the highest value in that range which is equal to the URE for nickel subsulfide. The low end of the range is equal to 50% of the URE for nickel subsulfide. The RfC value for nickel subsulfide was also used all nickel compounds (including nickel oxide).

[m] See Section 2.1.1.2 of the TSD for a description of the PAH/POM grouping.

[n] Although trichloroethylene is carcinogenic by a mutagenic mode of action, the age-dependent adjust factor for the URE only applies to the portion of the slope factor reflecting risk of kidney cancer. As such, the URE is adjusted by a factor of 1.12 (rather than the typical factor of 1.6).

[o] The individual xylene isomers were combined and noncancer impacts were estimated using the RfC for their mixture.

This page intentionally left blank.

Appendix I

Adjustments from the 2011 Emissions/Modeling Approach

In a small number of situations, tract modeling results were adjusted due to errors or anomalies that had impacts on the resultant tract risks.

For a few of these tracts, the modelled data were adjusted due to uncertainty with respect to the surrogate used to allocate the county-level off-network (parking) emissions to the tract. These few tracts were each in a highly urban core and were allocated based on square footage of industrial, commercial, institutional, and residential as an indicator of parked cars. However, the surrounding tracts appeared very similar with respect to parking areas. Given the uncertainty at this resolution, we chose to substitute the onroad light-duty values from a nearby tract. We chose the next-highest onroad light-duty risk and associated concentrations and exposures.

Exhibit I-1 contains the adjustments and the rationale.

Exhibit I-1. Adjustments to Tract-level data

County, State	Tract	HEM Run Groups	Pollut-ants	Data to be Adjusted	Adjustment	Reason
Kern, CA	06029001600	OR-LD OR-HD	all	Conc., exposures, poll.-specific risks (cancer and noncancer) by HEM run group and source group	Recompute as the county mean using all tracts in the county except for 06029001600	Faulty Surrogate 200 (urban primary road miles) puts 100% of the emissions in this tract, yet there appears to be no primary roads in the tract
Jefferson, IL	17081051000, 17081050900	OR-LD OR-HD	all	Same as above	Recompute as the county mean using all tracts in the county except for 1708105100 and 17081050900	Faulty Surrogate 200 (urban primary road miles) puts 100% of the emissions in these two tracts, yet there appears to be no primary roads
King, WA	53033007402, 53033007401	OR-HD NONROAD	all	Same as above	Recompute as the county mean using all tracts in the county except for 53033007402 and 53033007401	Unreasonable emissions density in high-population tracts from nonroad due to surrogates 140 and 100 (used for 520 due to gap fill issue) and from surrogate 221 in onroad. These two tracts were split up since last NATA.
San Francisco, CA	06075011700, 06075061500	NR-Gas/Other Cancer	all	Same as above	Recompute as the county mean using all tracts in the county except for 06075011700 and 06075061500	Surrogate 520 created large emission densities for commercial lawn and garden emissions that are likely not prevalent in this tract

County, State	Tract	HEM Run Groups	Pollut-ants	Data to be Adjusted	Adjustment	Reason
Du Page County, IL	17043843900	OR-HD_Diesel (source group)	all	Same as above	Recompute as the county mean using all tracts in the county except for 17043843900	The risk is due to surrogate 205 (truck stops). The underlying truck-stop data shapefile shows the weigh station to be no longer in operation. The satellite data do not show any truck stops in this tract.
St. Louis County, MO	29189221422	OR-HD_Diesel (source group)	all	Same as above	Recompute as the county mean using all tracts in the county except for 29189221422	The risk is due to surrogate 205 (truck stops). The underlying truck stop data shapefile shows the major truck stop that caused 71% of the county emissions to be allocated into this tract is actually in the neighboring tract to the east, and thus this tract should not have gotten the emissions.
Orange County, CA	06059075514	NR-Gas/Other Cancer Risk	all	same as above	Recompute as the county mean using all tracts in the county except for 06059075514	Very high emission densities caused by two surrogates: 520 (commercial + industrial + institutional) and 510 (commercial + industrial) used for allocating commercial lawn and garden equipment and commercial equipment that do not appear consistent with the land use in that tract
Lehigh County, PA	42077001000	OR-Light Duty	all	Total	Set onroad–light-duty risk to next-highest neighboring tract in the county (42077000800)	This is the only tract with risk > 100-in-1 million in an urban area due to onroad risk from parking-area emissions. It appears to have similar characteristics (with respect to parking areas) as neighboring tracts which are about 10-in-1 million or so lower risk.
Lancaster County, PA	42071000100	OR-Light Duty	all	Total	Set onroad–light-duty risk to next-highest neighboring tract in the county (42071000700)	Same as above
Hennepin, MN	27053104400	OR-Light Duty	all	Total	Set onroad–light-duty risk to next-highest neighboring tract in the county (27053126100)	Same as above

County, State	Tract	HEM Run Groups	Pollut-ants	Data to be Adjusted	Adjustment	Reason
Hamilton, OH	39061000700	OR-Light Duty	all	Total	Set onroad–light-duty risk to next-highest neighboring tract in the county (39061000900)	Same as above
Tulsa, OK	40143002500	OR-Light Duty	all	Total	Set onroad–light-duty risk to next-highest neighboring tract in the county (40143003300)	Same as above. Additionally, the primary road on the perimeter of the tract boundary may also be in partly-adjacent tract.
Puerto Rico: 2 Municipios: San Juan 72127 and Ponce 72113.	Tracts impacted by CMV	CMV - ports	all	Emissions, tract concentrations	Use same emissions and concentration as 2005 NATA, recompute risks based on the concentrations	2011 NEI did not have port emissions, so gap-fill with 2005 NATA