Sonoma Technology, Inc.

1360 Redwood Way, Suite C
Petaluma, CA 94954-1169
707/665-9900
FAX 707/665-9800
www.sonomatech.com

ANALYSIS OF AIR TOXICS MONITORING DATA WORK PLAN

STI-903555-2442-WP2

Prepared by:
Hilary R. Hafner
Neil J. M. Wheeler
Paul T. Roberts
Sonoma Technology, Inc.
1360 Redwood Way, Suite C
Petaluma, CA 94954-1169

Prepared for:
Lake Michigan Air Directors Consortium
2250 East Devon Avenue
Des Plaines, IL 60018

January 23, 2004

This page is intentionally blank.

TABLE OF CONTENTS

LIST OF TABLES

1. INTRODUCTION

This project consists of the analysis of air toxics monitoring data and evaluation of air toxics modeling. Both facets of the project build upon previous and ongoing analyses and evaluations. This introductory section provides an overview of the project objectives, a summary of previous analyses of air toxics data, and an overview of the proposed data analysis efforts. The proposed modeling efforts are discussed in Section 3.

1.1 OBJECTIVES

Current Government Performance and Results Act (GPRA)[1] commitments "specify a goal of reducing air toxics emissions by 75% from 1993 levels to significantly reduce the risk to Americans of cancer and other serious adverse health effects caused by airborne toxics (Ryerson et al., 2001)." The national air toxics program has been designed with a multi-faceted approach, combining many agencies' activities, to provide the information needed to address and eventually meet the GPRA goal. One of the activities is the National Air Toxics Assessment (NATA), which includes air toxics monitoring made at regional, community, and neighborhood/micro scales and subsequent data analyses. On a national or regional scale, measurements and analyses are made to investigate spatial and temporal trends, establish background concentrations, and perform comparisons among regions. On a neighborhood/micro scale, measurements are made to investigate specific compounds of local concern, investigate specific sources or source types, and generally focus on small areas with potentially high air toxics concentrations. Community-scale measurements focus on providing data to characterize typical community exposure and to assess the associated risk.

The ultimate objective of an analysis of community-scale air toxics monitoring data is to be able to answer a range of policy-relevant questions related to community exposure. As part of the planning for this project, a list of policy-relevant questions was compiled and sent to stakeholders[2] for review and comment. The resulting list is presented in Section 2.4 and was used to develop analyses that are policy relevant. Inspection of the list reveals a wide range of policy-relevant questions; to address them all comprehensively is beyond the scope of this current work.

The objectives of this project are as follows:

- Provide a comprehensive "look" at all of the existing air toxics data—both the historical archive and the pilot city database. This "look" will provide a broad national assessment of air toxics data, and a detailed local examination in a few select areas. The analyses will go beyond the initial studies by Battelle and STI (Battelle Memorial Institute and Sonoma Technology, 2003; Bortnick et al., 2001).

[1] The Government Performance and Results Act of 1993 (GPRA) requires that federally funded agencies develop and implement an accountability system based on performance measurement, including setting goals and objectives and measuring progress toward achieving them.

[2] Stakeholders are identified as the state air toxics monitoring community and local, state, and federal air toxics policymakers.

- Present a clear message for policy-makers about air toxics concentrations across the country from both national-level and community-level perspectives.
- Provide guidance and tools that will enable agencies collecting air toxics data to look at and use their data.
- Support the U.S. Environmental Protection Agency's (EPA) new air toxics data analysis web site.
- Prepare limited model-to-monitor comparisons, which complement the EPA's ongoing modeling work.

1.2 SUMMARY OF PREVIOUS AIR TOXICS ANALYSIS

A wealth of air toxics monitoring data has been collected by state and local agencies in support of various monitoring programs over several decades, including the recent 10-city pilot monitoring project (Nizich, 2001). Several recent data analyses focused on providing recommendations for the design of the national air toxics monitoring network including recommending the number of national trends sites, the number of monitors per city, the spatial distribution of monitors (location), the frequency of sample collection, the species to monitor, sample duration guidelines, and the handling of data below minimum detection limits (MDLs[3]) in analyses (Battelle Memorial Institute and Sonoma Technology, 2003; Bortnick et al., 2001; Brown and Hafner, 2003; Brown et al., 2003; Hafner and Brown, 2003b; Hafner et al., 2003).

Specific analyses in support of national network design performed as part of the referenced work included:

- Spatial analysis (i.e., analyses to assess the placement of monitors to best measure the impact of diesel particulate matter [DPM] on the population).
- Analysis of the variability of organic carbon (OC) and elemental carbon (EC) concentrations across the United States (e.g., are diesel emissions as important to EC concentrations across the United States compared to Los Angeles?). These analyses demonstrated the usefulness of data from other programs to enhance and improve analyses (e.g., Photochemical Assessment Monitoring Stations [PAMS], fine particulate matter [$PM_{2.5}$] speciation network).
- Source apportionment using air toxics and $PM_{2.5}$ data (e.g., which air toxics program and supplemental data are useful in apportioning DPM?). These analyses showed the need to work toward lower MDLs for most toxics species.
- Case-study analysis on the relationship between meteorology and toxic pollutant concentrations. These analyses indicated the need for collocated surface meteorology with air toxics measurements.

[3] Also referred to as method detection limit

1.3 OVERVIEW OF AIR TOXICS DATA ANALYSIS

Analyses of air toxics data need to be performed to address the monitoring objectives. In the Air Toxics Monitoring Concept Paper (U.S. Environmental Protection Agency, 2000), three objectives for the national monitoring program were identified: (1) establish trends and evaluate effectiveness of hazardous air pollutant (HAP) reduction strategies at the national level, (2) characterize ambient concentrations (and deposition) in local communities, and (3) provide data to support and evaluate dispersion and deposition models). Objectives for state and local air toxics monitoring programs include these same three objectives, as well as to investigate source-specific (compliance-related) issues and to support risk assessments.

In this study, STI will go beyond recent air toxics data analyses, which focused on providing network design recommendations, to characterize air toxics concentrations at both the national and local (urban) scales. To ensure that these results are policy-relevant, the analyses will address three general questions:

1. How good are the data?
2. What are the air toxics concentration levels nationally and locally?
3. What do air toxics data say about the effectiveness of various control programs?

Graphical depictions of the data will be used to explore temporal and spatial variability in species concentrations and relationships, with the goal of identifying potential sources and the exposure risk to the population. The data will be used to develop guidance for stakeholders on data validation procedures and screening criteria, and recommended analyses.

To address how the data can be used by the stakeholders, the EPA has proposed a series of prescribed analyses (**Table 1-1**). Analyses will be made for selected communities to demonstrate their data needs and to illustrate the utility of the analyses to answer policy-relevant questions.

Table 1-1. Proposed prescribed questions and example analyses for toxics data (Hafner, 2003a).

Questions	Example Analyses
Are data of sufficient quality for analysis?	
How have the data been validated? Are there any years or species measurements that have suspect data quality?	Run screening checks; identify outliers. Review summary statistics; identify data ranges.
Are there sufficient samples available for detailed analyses?	Determine the number of samples/species with concentrations above detection.
How would you characterize toxics in the area?	
Which are the five most abundant toxics species at each site (both on a concentration and risk-weighted basis)?	Determine the median concentrations, concentration ranges, and number of samples; rank the species.
How do these species vary by season, month, (and time of day, if available)?	Prepare box plots of concentrations by month (and time of day) for abundant species; inspect patterns.
Do these abundant species show any day-of-week patterns?	Prepare time series plots of these species and box plots by day of week; inspect patterns.
What are the levels of toxics species for high-concentration days vs. low-concentration days?	Compare box plots of species concentrations and median fingerprints for high- vs. low-concentration days; investigate differences.
What are the local sources of toxics?	
What are the potential toxics sources in the area?	Prepare a map showing 20 major VOC and/or $PM_{2.5}$ sources; compare emissions to monitor site placement. Investigate satellite photos for the area; identify point, mobile, and area sources near the monitor.
Do the toxics data corroborate the source mixture?	Examine key species noted as tracers and prepare scatter plots and correlation matrices; determine species relationships. Prepare box plots by day of week; investigate weekday/weekend trends.
How does meteorology affect toxics concentrations?	
What are the meteorological conditions on high-concentration days vs. low-concentration days?	Prepare scatter plots of meteorological variables vs. toxics concentrations and fingerprint plots comparing high- and low-concentration days.
What species are likely transported into the area and from where?	Prepare wind and pollution roses of high- and low-concentration days; investigate differences.
Can upper-air meteorology be used to corroborate transport?	Prepare trajectories, wind runs, and/or ventilation analyses at selected heights.

Table 1-1. Proposed prescribed questions and example analyses for toxics data (Hafner, 2003a).

Questions	Example Analyses
Are there changes in toxics concentrations over time?	
What are the annual trends in toxics species?	Prepare annual box plots of key species (including benzene) and top five most abundant species; evaluate trends. Identify controls implemented during the data record; assess whether toxics trends might be in response to controls.
What changes have occurred in motor vehicle emission species?	Inspect box plots of key species (e.g., benzene, formaldehyde) by year; assess potential declines in ambient benzene with implementation of reformulated gasoline.
What changes have occurred in other major sources in the area?	Prepare box plots of key species and median fingerprints by year; assess the data for trends attributable to major sources.
What changes have occurred in area sources?	Prepare box plots of key species and median fingerprints by year; assess the data for trends attributable to area sources.

1.4 OVERVIEW OF MODELING PLANS

Air quality models are mathematical tools that simulate atmospheric processes, including transport, diffusion, transformation, and removal. Air quality models have been used for decades to assess the potential impact of emission sources on ambient concentrations of criteria and toxic air pollutants. In the past decade, air quality models have also been used as planning tools for criteria pollutants. However, until recently, air quality models have not been used as planning tools for toxics, due to the lack of measurements with which to evaluate the models. There is a clear need to assess the usefulness of these models in air quality planning and to improve both modeling and evaluation methods. A recent review of air toxics modeling was prepared under contract to the Coordinating Research Council (Seigneur et al., 2002b). As recommended by that review, model simulation results should be compared with available data to determine their reliability. In fact, one of the objectives of the national monitoring program is to "provide data to support and evaluate dispersion and deposition models" (Nizich, 2001). The purpose of this activity is to evaluate the performance of one or more air quality models using data from two pilot cities.

In this study, STI will explore model performance issues using measurement data from the 10-city pilot. The modeling and evaluation will focus on the cities of Detroit and Seattle. The presence of large water bodies adjacent to these cities produces complex meteorological conditions that will significantly challenge the models being evaluated.

Because toxics are assessed using annual concentrations, air quality models for toxics must be applied over entire years to obtain annual concentrations or the annual concentrations derived from a composite of numerous shorter periods. As a result, air quality models for toxics

have used simpler meteorological inputs and simpler (if any) representations of chemical transformations than those used in planning for criteria pollutants (e.g., ozone). These simplifications can lead to errors in both the pollutant concentration and spatial distribution, which reduces a model's utility as a planning tool. Through the model evaluation process, we will identify potential modeling issues that will guide our selection of alternative models to apply and evaluate.

1.5 CONTENTS OF THE WORK PLAN

The work plan consists of the technical approach for the characterization of air toxics (Section 2), the technical approach for the model evaluation (Section 3), the schedule and deliverables (Section 4), management plan (Section 5), and references (Section 6). Appendix A presents the policy questions that were sent to stakeholders for review in September 2003. Appendix B shows the list of toxics modeling discussion items sent to stakeholders for comments in September 2003.

2. TECHNICAL APPROACH – CHARACTERIZATION OF AIR TOXICS

2.1 OVERVIEW

A two-phase approach to the air toxics characterization portion of the project was requested by the awarding committee. In the first phase, we will obtain an overall appreciation (i.e., national review) of the data in the archive of both historical and pilot-city data; this analysis will build upon the previous analyses and will focus on community-level concerns (e.g., to what level of air toxics is my community being exposed?). This phase will also include data validation. In the second phase, we will perform more focused analyses of data from selected areas of the country (i.e., prescribed analyses). These analyses will be tailored to the individual areas to account for previous data analyses and the current understanding of air toxics in those areas. The steps in this two-phase approach include:

1. *Acquire, review, and validate the data.* As a part of the data validation, develop guidelines for the states and address questions developed from stakeholder feedback (see questions at the beginning of Section 2.3).
2. Provide guidance for web site development by the Cooperative Institute for Research in the Atmosphere (CIRA) at Colorado State University for the EPA. Review data displays, data handling (e.g., how to treat data below the MDL, how to create city-wide averages), and data accessibility.
3. *Characterize the air toxics data.* Identify species at key sites that have adequate data for meaningful analyses. Prepare data displays, focusing on inter-site comparisons within a community, areas of different land use and emissions density, temporal variations, and periods of known differences or changes in measurement or analytical technique.
4. *Interpret and further analyze the data.* Include focused analysis on data from cities with a number of sites (i.e., three or more) with adequate data. These analyses will include meteorological analyses, emission-inventory/spatial-mapping analyses, and source apportionment analyses, all of which are dependent on the amount, type, and availability of air toxics and supporting data. Policy relevance of analyses will then be assessed (see questions at the beginning of Section 2.4).

The air toxics characterization portion of this project consists of the following tasks, detailed in the next sections.

- Task 1: Prepare a work plan
- Task 2: Acquire, review, and validate the data
- Task 3: Describe, display, analyze and interpret air toxics data
- Task 4: Report and synthesize

2.2 TASK 1: PREPARE A WORK PLAN

This work plan was prepared to meet the requirements of the project. The first step in preparing this work plan was to obtain feedback from stakeholders regarding the types of policy questions upon which analyses should be focused. A list of issues was developed by STI and LADCO and sent to stakeholders for comment (see Appendix A); results from the questions in

the issues list were used to develop the technical approach discussed in this work plan. Data quality questions are summarized in Section 2.3 and policy-relevant questions are summarized in Section 2.4.

2.3 TASK 2: ACQUIRE, REVIEW, AND VALIDATE THE DATA

The historical and pilot city toxics database analyses were conducted primarily with Level 1[4] quality control (QC) data (i.e., extensive work was performed by Battelle to check the completeness of records, document methods, document MDLs, etc.) and less contractual focus was placed on the quality or internal consistency of the data (e.g., Level 2[5] validation). However, at recent toxics data validation training sessions for EPA (Hafner, 2003a), state and regional stakeholders expressed the need for the development of a validation approach and customized screening criteria, as well as continued validation tool development similar to previous efforts for the PAMS speciated hydrocarbon, $PM_{2.5}$, and air toxics data (e.g., Brown and Main, 2002; Main and Collins, 1994; Main, 2000; Main and O'Brien, 2001; Main, 2002; Main and Brown, 2002a, b). It is important to future toxics trends network and community-level analyses that guidelines are established for data validation.

The questions for Task 2 that help identify whether a community's data can be used to address policy-related questions are outlined next. Task 2 analyses will be linked to these questions.

Questions to Identify Whether a Community's Data can be used to Address Policy-Related Questions

- Which toxics species are adequately represented in the database?
- Can policy questions be addressed with the available and planned data?
- How should missing data and data below detection levels be treated, and how do different data treatments affect data analysis and modeling results?
- How good are the historical and pilot city data and what are the criteria for determining "good"?
- What is our confidence in the data?
- Are data useful even if they don't meet QA/QC and data completeness criteria?

2.3.1 Data Acquisition and Preparation

STI will acquire two sets of air toxics data from the EPA and Battelle: one historical compilation of more than 30 years of air toxics information for some sites and pollutants, and a 10-city pilot study data set consisting of a year's worth of data. The historical database will be obtained from EPA and documentation provided by Battelle (e.g., Bortnick et al., 2001). The

[4] Level 1 validation consists of routine checks during the initial data processing and generation of data including proper data file identification; review of unusual events, field data sheets, and result reports; and checks of instrument performance checks and deterministic relationships.

[5] Level 2 validation consists of internal consistency tests to identify values in the data that appear atypical when compared to values of the entire data set.

pilot-city database and summary statistics will be obtained from Battelle. We also will coordinate with the EPA and CIRA regarding data sets; CIRA is part of an EPA-sponsored project to place the toxics data on an interactive web site.

One of the necessary steps in preparing the data set for validation and analysis is to determine a treatment scheme for data below the MDL. Considerable work on this issue has been performed in the previous analyses by Battelle. The selection of a treatment scheme is a function of the analysis objective. Previous work showed that for most analyses, substituting MDL/2 for data below detection (listed as zero or "non-detect"), and retaining data reported with a value below the MDL, was a reasonable choice. We propose to employ this substitution method for most of our analyses unless the analysis objective requires a different approach. In addition, we intend to build on the previous work by providing new flags in the database that document the percent of data values below the MDL used to produce the annual and seasonal averages.

2.3.2 Data Review

Species Selection. Given the size of the historical database, we will not be able to investigate every compound. However, focusing on a modest number of air toxics species can illustrate an approach to data validation and analysis that can later be applied to other air toxics. There are currently 188 HAPs or air toxics regulated under the Clean Air Act (CAA). Of these, a list of 33 urban air toxics was developed by EPA; these toxics are thought to have the greatest impact on the public and the environment in urban areas. From the urban list, 18 air toxics were identified as core to the 10-city pilot study. Of these, six air toxics are identified as the main National Air Toxics Trends Sites (NATTS) compounds. **Table 2-1** lists the 33, 18, and 6 important air toxics. Further complicating species selection, the metals data are reported in coarse, fine, PM_{10}, and total suspected particulate (TSP) size selections. Most of our analyses will focus on the 18 air toxics species/species groups.

Table 2-1. EPA's list of 33 urban air toxics; the core 18 air toxics are shown in bold text, and the seven main NATTS compounds are denoted with an asterisk.

acetaldehyde	diesel particulate matter*****	mercury compounds
acrolein*	dioxin	**methylene chloride**
acrylonitrile	ethylene dibromide	**nickel compounds**
arsenic compounds*	**propylene dichloride**	polychlorinated biphenyls (PCBs)
benzene*	1, 3-dichloropropene	polycyclic organic matter (POM)
beryllium compounds	ethylene dichloride	quinoline
1, 3-butadiene*	ethylene oxide	1, 1, 2, 2-tetrachloroethane
cadmium compounds	**formaldehyde***	**perchloroethylene**
carbon tetrachloride	hexachlorobenzene	**trichloroethylene**
chloroform	hydrazine	**vinyl chloride**
chromium compounds (hexavalent chromium*)	**lead compounds**	
coke oven emissions	**manganese compounds**	

Database Selection. Two approaches to database selection were considered: either first explore the historical database or first explore the pilot-city database. There are advantages and disadvantages to both approaches (e.g., the pilot-city data were collected with the same measurement methods, or at least the differences were known; the historical data record is longer and contains more sites, but has more "problems" [Bortnick et al., 2001]). Given the recent focus on the pilot-city data, and in response to concerns that the historical data have not been as fully "mined" as they could be, we will start with the historical data set for the development of data validation guidelines and prescribed analyses. The "lessons learned" from the historical database will then be applied to pilot-city data. The focus of the analyses will be on recent trends and current levels of air toxics from the past 10 years (1993 to 2002); however, we will investigate the entire data record for selected pollutants and sites. With the more current years of data, some of the issues present in earlier data's sampling and analytical methodologies are avoided.

Site Selection. We will assess data completeness (i.e., samples reported vs. samples expected, regardless of concentration above or below the MDL) for the selected species by site and city to determine which sites and cities have an adequate amount of data for useful analysis in later tasks. We will also assess how many of the samples were below detection for the target species. We propose a lower-end cut of 50% completeness, regardless of whether the concentrations were below detection, for each specie during the period of interest. This type of screening allows retention of useful periods of data even though sampling may not have occurred throughout a 10-year period, for example. We will prepare summary statistics for sites, species, and years to assess the ranges of concentrations and variability. Both data completeness and summary statistics will be built upon previous computations.

Sites/cities for further review will be selected based on data availability, geographical distribution, abundance (or lack) of previous analyses, emission sources, and other characteristics (e.g., meteorology, proximity to the Canadian or Mexican border, topography, etc.) Other data acquisition and site selection considerations include the number of species available at a single site, temporal resolution of the data, availability of supporting data (e.g., criteria pollutants, surface meteorology), changes in sampling and/or analysis methodology (some of this information is available in Bortnick et al., 2001), and changes in local emissions (stakeholder input will be useful in this regard). A priority list of cities (some cities have multiple toxics sites) will be prepared and discussed with the stakeholders. From this list, we will select 15 to 25 cities for further analyses. For those cities identified as having a useful data set, we will also acquire selected supplemental data, such as PAMS, continuous PM mass, criteria pollutant, speciated PM (e.g., organic and elemental carbon, sulfate, nitrate) and meteorological data.

2.3.3 Data Validation

The goals of the data validation step are to prepare the data for further analysis and to develop guidance for the states in validating their data collected as a part of their air toxics monitoring programs. STI will apply a data validation approach used in previous projects for volatile organic compounds (VOC), $PM_{2.5}$, toxics, and DRUM[6] speciated PM data (e.g., Brown

[6] Davis Rotating-drum Universal-size-cut Monitoring.

and Main, 2002; Hafner and Brown, 2003a; Main, 2002; Ryan et al., 2003). The air toxics validation checks to be performed will follow guidelines set forth in the PAMS and $PM_{2.5}$ data analysis workbooks (Main and Roberts, 2000, 2001), previous toxics analyses (e.g., Hafner et al., 2002; Main, 2002), and previous data validation training (Hafner, 2003b). Checks include range checks, inter-species correlations, inter-site correlations, and graphical review of samples (e.g., time series, scatter, and fingerprint plots).

Battelle performed extensive data quality investigations such as checking validity flags, ensuring that date/time/site labels were correct, attaching method codes and units to concentrations, and ensuring that MDL values were provided for each concentration. Battelle also spent considerable effort to remove duplicate records, create 24-hr averages, standardize units, and apply completeness tests prior to their analyses. All of these tasks served to clean up and document the database for future use. Summaries of the data above the detection limit were also prepared for 18 HAPs. We will work from this data set and perform the following additional data validation steps:

- Prepare scatter plots, time series plots, and other graphic methods to investigate the selected data and recommend the types of plots that proved most useful (and what to look for in the data on the plots).
- Prepare and inspect summary statistics including minimum, maximum, median, mean, standard deviation, and coefficient of variation for species by site and year. The summary statistics are also used to assess data completeness, provide range checks, provide a comparison to MDLs, and to determine the most abundant air toxics on a concentration or risk-weighted basis.
- Investigate internal consistency checks. For example, check that the sum of identified hydrocarbon or PM species concentrations in a sample are less than the total hydrocarbon or PM mass reported for the sample. These types of checks will require data in addition to the 33 HAPs and may not be readily available. Checks from which to start include those outlined in the workbooks and in a recently developed data validation workbook (Hafner, 2003).
- Develop general screening criteria based on chemical fundamentals (e.g., particle size, atmospheric half life) and relationships observed in the data (e.g., ratios, range checks) to facilitate data validation and improve data quality. For example, if species are emitted by the same source, their ambient concentrations should correlate well if the samples are dominated by that source. Analyses may lead to the development of methods to customize these criteria for different site types or parts of the country.
- Use the newly enhanced capabilities of VOCDat to graphically review the toxics data—this includes using toxicity and cancer-risk factors to scale species, similar to using scales for ozone reactivity. Assess the utility of this approach.
- Inspect and flag the selected data. Flagged species will be documented with the reasoning behind the flags.

Tools for validation include VOCDat software (Main and Prouty, 2000) and commercially available software, such as SYSTAT statistical software, to visually inspect data and prepare summary statistics. VOCDat was recently updated through EPA funding to enhance

its use for air toxics data validation, including upgraded import/export routines, expansion of species lists, and use of toxicity and cancer-risk weighting of the data to aid in the investigation of exposure (e.g., Hafner, 2003a). We have included some programming time in the budget of this task to make several small enhancements to VOCDat, if necessary, to facilitate the data validation process. Stakeholders identified the need for additional tool development for use in validation and analysis of toxics data. For example, one of the deliverables for this task will be example screening criteria for toxics to be used by VOCDat.

The outcome of the data validation will be (1) validated data for 18 or more HAPs at up to 10 cities over the latest 10 years (historical database), (2) additional database flags useful to data analysts, and (3) data validation guidelines (i.e., example screening criteria and checks, types of plots to explore, what to look for in the plots and screening checks, reasonable concentration ranges, etc.) based on our findings.

2.3.4 Database Development Guidance

We will provide external review to CIRA on their EPA-sponsored web-based database development project. We will review the web site and provide CIRA with recommended default decisions regarding MDL substitution, averaging requirements (daily, annual, city-wide), data completeness, etc. and also recommend the selection of other options. The deliverables will be e-mail correspondence and conference call discussion regarding the web database design.

2.4 TASK 3: DESCRIBE, DISPLAY, ANALYZE AND INTERPRET AIR TOXICS DATA

The objective of this task is to develop ways to use the air toxics data to address policy-relevant questions such as “What are the air toxics concentrations nationally and locally?” and “What do air toxics data tell us about the effectiveness of emission controls?”. A common theme in the detailed list of questions provided next is the need to understand, at a community level, to what levels of air toxics people are exposed. For example, cities want to understand how air toxics concentrations vary within the city, compare to other cities, compare to regional background concentrations, and have changed over time, and how such data can best be used to better understand exposure. This task will begin to address some of these questions using example data sets.

Policy-relevant Questions to be Addressed with Air Toxics Data

How do air toxics concentrations vary nationally and locally?

- Geographic variation: What does a broad national assessment say about air toxics concentrations across the country?
- City-to-city and intra-urban variation: Based on case studies, what can we say about air toxics concentrations on the urban scale? How do urban area concentrations compare with those in nearby rural areas? Are there typical urban and rural profiles (fingerprints)? How does exposure to mobile source toxics vary as a function of distance from roadways? How can concentrations from a small network be extrapolated to other areas? How representative are the existing monitoring sites (i.e., how broadly [spatially] can the annual average concentration at a particular site be applied)? Is more detailed "standard metadata" needed to better define the specific, micro-scale characteristics of air toxics monitoring sites?
- Baseline (background) concentrations: What is meant by "background" levels? How can background levels be estimated? What are reasonable estimates of background levels?

How do air toxics concentrations vary temporally?

- What can we say about the variation in air toxics concentrations on yearly (trends), seasonal (or monthly), day-of-week, and hour-of-day bases?
- How do we characterize "spikes" in ambient concentrations?
- What have we learned about the frequency of air toxics monitoring in order to capture the true story on exposure?
- What are appropriate methods for producing daily, annual, or city-wide averages?
- How do we quantify uncertainty in the data analysis results?

What do air toxics data tell us about the effectiveness of emission controls?

- How effective have mobile source controls been in reducing exposure to toxics?
- How effective have maximum achievable control technology (MACT) standards been in reducing exposure to toxics?
- What effect have various community assessment projects had in reducing exposure to air toxics?
- Have emissions controls reduced background concentrations?
- Should other non-toxic species measurements be added to toxics monitoring sites to aid source apportionment?
- How has implementation of ozone and PM controls reduced air toxics levels (and vice versa)?

2.4.1 Describe and Display

Once the data have been validated and screened for completeness, the next step is to describe and display the data in order to better understand spatial and temporal variations in air toxics. Descriptions of the data will focus on inter-site comparisons within a community, areas of different land use and emissions density, temporal variations, and periods of known differences or changes in measurement or analytical technique. To enhance data analyses and interpretation, we will also provide additional information to the database including matching

site locations to the appropriate Metropolitan Statistical Areas (MSAs) and, for selected areas, provide site characteristics using geographical information services (GIS) information (e.g., population density, land-use type). Previous investigations by site classification code (i.e., urban, rural, commercial, industrial, mobile source) were inconclusive. Perhaps by augmenting the broad codes with more detailed information, differences among air toxics at various site types will be more clearly explained.

Data displays will build upon previously prepared displays and include new depictions of comparisons of toxics concentrations among regions of the country, sites within a city, urban vs. rural sites, different land-use types, varying emissions densities, and different seasons to obtain an overall understanding of national air toxics levels. Displays and summaries will include:

- Notched box plots. These plots are useful for displaying the central tendencies of species concentrations and exposure risk-weighted data (toxicity and cancer risk) both spatially (e.g., by site in a community or region, land use, emissions density, etc.) and temporally (e.g., by day of week, season, year, and time of day where available). These plots show that statistically significant differences between the quantities being compared exist when there is no overlap between the medians, plus/minus the confidence interval.
- Scatter plots and correlation matrices. These will be used to find relationships between species and sites, which can be further investigated by year, season, or day of week if enough data are available.
- Fingerprint plots. These plots show the composition of a sample (i.e., the concentration or weight percent of all the species identified in the sample). Average fingerprints will be investigated by year, season, time of day and day of week.
- Maps. These show valid annual average concentrations at all sites in the database for a given year and will, for example, help illustrate the national spatial distribution of monitoring sites and concentrations.

2.4.2 Analysis and Interpretation

In this task, data displays will be further interpreted, and more focused analyses will be made including national and MSA-level investigations of the spatial and temporal trends in air toxics concentrations, meteorological case study analyses, emission inventory mapping analyses, and source apportionment for selected areas similar to those preliminary analyses already performed for Detroit in the 2002/2003 Battelle/STI work (Battelle Memorial Institute and Sonoma Technology, 2003). These analyses will focus on species with adequate data completeness and quality, and on cities with a significant amount of data (i.e., many species, more than three sites), other supporting analyses (e.g., special studies), and the availability of meteorological data. These analyses also serve as examples of possible prescribed air toxics analyses.

Basic Analyses

The analyses will employ an approach similar to that applied to PAMS data (see Table 1-1), in which a list of suggested analyses are organized by a broad policy-related question

(are there changes in toxics concentrations over time?), analysis objectives (what changes have occurred in motor vehicle emission species?), and example analyses (for urban sites, inspect the central tendencies of motor vehicle emission tracer species by year and compare to controls that have been implemented and/or to fuel-composition data). The prescribed analyses range from a summary of the central tendencies of the data (e.g., annual averages of air toxics) to more in-depth case studies including source apportionment. For the selected areas we will perform prescribed analyses to provide data analysis and interpretation examples for stakeholders to follow in future investigations of their own data. We recognize the varying levels of stakeholder understanding of air toxics, of previous analyses, and of available data; thus, the focused analyses must be tailored for each selected area. The focus of the analyses will be on recent trends and current levels of air toxics (i.e., the past 10 years, 1993 to 2002); however, we will investigate the entire data record for selected pollutants and sites. In the more current years of data, some of the issues present in earlier data's sampling and analytical methodologies are avoided (e.g., significant differences in sampling and analytical methods). The draft prescribed analyses for air toxics will be used to more fully develop a suite of analyses that address specific questions and concepts identified by the stakeholders.

Focused Analyses

Depending on stakeholder needs, data availability, and the available budget, a selected number (i.e., five to ten) of more focused case studies will be conducted including spatial mapping of population density and emissions, identification of major meteorological influences among sites within a city or region, and source apportionment using positive matrix factorization (PMF). Other data, such as PAMS and $PM_{2.5}$ from the speciation network and continuous instruments, will also be used to augment the case-study analyses. These data often provide better spatial (additional sites) and temporal (less-than-24-hr samples) coverage than the toxics data. A number of different approaches will be used to find consensus, so that results are not based on only one type of graphical or statistical analysis, but are reached by a thorough understanding of the data. The more in-depth analyses (also called case studies) also provide good example analysis suites for other analysts. Example focused analyses include the following:

Spatial Analysis. The technical approach for this analysis is to use geographic information system (GIS)-based suitability modeling to assess potential locations for placing monitors to meet selected monitoring objectives. For example, in previous work, STI applied suitability modeling to assess diesel emissions in Detroit. The Environmental Systems Research Institute (ESRI) ArcGIS software, Spatial Analyst, will be used to perform the analyses for this study. ArcGIS provides a framework for overlaying and analyzing spatial relationships among multiple map layers. Spatial analysis techniques are used to investigate relationships among geographic features and sources of toxics emissions with respect to the existing toxics monitoring network. The Detroit analysis illustrated a technique for locating monitors to investigate diesel emissions impacts on the surrounding population. ArcGIS is a raster-, or grid-, based software package that provides a platform for working with gridded data sets. Spatial analysis is used to produce suitability maps highlighting "suitable" geographic regions derived from weighted and combined map layers based on established criteria.

Meteorological Case Studies. The objective of this analysis is to determine the effects of meteorology on toxics concentrations using a case-study approach. Case studies are in-depth analyses of selected time periods with interesting characteristics (e.g., high concentrations of a particular pollutant). In previous work, the Detroit metropolitan area was selected to coordinate with other in-depth investigations of Detroit data; this case study is illustrative of the types of analyses we propose to perform for selected cities. General meteorological patterns were determined by looking at the U.S. Department of Commerce/National Oceanic and Atmospheric Administration (NOAA) Daily Weather Maps and NOAA HYSPLIT ETA Data Assimilation System (EDAS) model back trajectories. Local flows were analyzed using surface meteorological wind data available at stations near toxics monitoring stations. Days with high concentrations of chemical species/toxics at one site, but not at other sites within the region, were identified for use in case-study analyses. Synoptic-scale weather systems played an important role in transport of toxics from the Detroit urban area to more rural regions. In addition, high toxics concentrations were observed in rural regions as part of an air mass that did not originate in Detroit. Mesoscale flows in the Detroit region were found to be common as a result of Detroit's proximity to Lakes Huron and Erie, particularly when high pressure resided over the region. While general relationships could be investigated, it was difficult to determine more specific effects of meteorology on toxics concentrations due to the lack of continuous, collocated toxics measurements and the lack of upper-air wind and temperature data.

Multivariate Receptor Modeling. Receptor modeling is a mathematical procedure for identifying and quantifying the sources of ambient air contaminants at a receptor, primarily on the basis of ambient concentration measurements at that receptor (also called source apportionment). Multivariate receptor models require the input of data from multiple samples and extract the source apportionment information from all of the sample data simultaneously. The reward for the extra complexity of these models is that they estimate not only the source contributions but also the source compositions (profiles).

In a case study for a city, we will first (1) examine and interpret species relationships at individual sites and evaluate how these relationships change between sites; (2) inspect temporal differences between sites; and (3) perform preliminary factor analysis to further investigate species relationships and identify differences among sites within a city. We will then perform exploratory source apportionment to explore possible emission source types. PMF will be used to identify likely sources. The factor analysis tools provide factors that can be related to emission source types and estimate the quantitative contribution of each factor in every sample. Thus, the variation of source strength by time of day and wind direction (if data are sufficiently resolved), day of week, and season can be explored.

The deliverables for this task include a suite of prescribed analyses for air toxics that is scalable based on data availability and example prescribed analyses for the selected cities.

2.5 TASK 4: REPORTING AND SYNTHESIS

Throughout the project, we will strive to prepare documents (e.g., reports, executive summaries) that are usable by readers with a range of backgrounds and knowledge regarding air toxics. Many of the policy issues described for Tasks 2 and 3 have been addressed in previous

analyses. The previous reports were organized to answer the network design questions. One of the first tasks in reporting and synthesis will be to summarize the previous results with respect to the current list of policy-related questions in a technical memorandum.

The major focus of this task is to synthesize the results from data validation and data analysis—including inter-site comparisons, land-use comparisons, temporal and seasonal trends, exposure risks, and source apportionment—in documents that clearly communicate key findings and recommendations. Deliverables to be considered through discussion with LADCO include technical memoranda, technical reports with executive summaries, "white papers", presentations, newsletters, outline and ideas for a Toxics Data Analysis Workbook, and brochures or flyers. An example of a key deliverable for the stakeholders is a technical report illustrating a case study of a selected city including specific findings and recommendations related to the policy questions outlined earlier. In this way, the selected cities can use their results as the basis for future analysis. The results will also be clearly summarized and explained in a PowerPoint presentation at a follow-up meeting designated by LADCO in early/mid-2004.

3. TECHNICAL APPROACH – MODEL EVALUATION

3.1 INTRODUCTION

Until recently, air quality models have not been used as planning tools for toxics controls, due to the lack of measurements with which to evaluate the models. There is a clear need to assess the usefulness of these models in air quality planning and to improve both modeling and evaluation methods. One of the objectives of the national toxics monitoring program is to provide data to support and evaluate dispersion and deposition models.

There are a number of broad policy-relevant questions that can be addressed through modeling analyses, such as

- Are the models good enough for toxics assessments?
- Are the models good enough to use in air quality planning?
- What are the uncertainties in modeling?
- How should models be used in risk assessments?

Each of these broad questions can be further divided into numerous specific technical questions. It is not the intent of this study to address all of these questions. Instead we will focus on one technical issue that, in part, answers both how good models are and what the uncertainties are in modeling:

> *How do model predictions, relative to observations, improve as the representation of meteorology increases in detail?*

3.2 OVERVIEW

The Assessment System for Population Exposure Nationwide (ASPEN) (Rosenbaum et al., 1999) was used to model toxics during the National Air Toxics Assessment (NATA). Evaluations of the ASPEN dispersion module (ASPENA) following the NATA found that, in general, at the exact monitor locations, the model estimates were lower than the monitor averages for most of the pollutant/monitor combinations (U.S. Environmental Protection Agency, 2003). For urban-scale toxics modeling, the Industrial Source Complex Short Term model version 3 (ISCST3) (U.S. Environmental Protection Agency, 1995) is recommended in EPA guidance on urban-scale toxics modeling (U.S. Environmental Protection Agency, 1999a).

In this study, STI will explore model performance issues using measurement data from the 10-city pilot. The modeling and evaluation will focus on the cities of Detroit and Seattle. The presence of large water bodies adjacent to these cities produces complex meteorological conditions that will significantly challenge the models being evaluated.

ISCST3 will be applied initially for the periods that toxics data were collected in these cities, using representative meteorological data and updated emissions inventories. This will allow us to compare the results of this study with modeling results from the NATA national-scale assessment, and potentially extend the conclusions of the model evaluation to other areas.

After the initial modeling and evaluation, STI will prepare non-steady-state wind fields with a diagnostic meteorological model, apply the Comprehensive Air Quality Model with Extensions (CAMx) (ENVIRON, 2003) at the same resolution as ISCST3, and evaluate the model predictions. The purpose of this additional modeling is to determine whether model performance issues identified in the initial evaluation can be addressed by more detailed treatment of the meteorology in the modeling system.

The results of the alternative modeling will be evaluated using the same methods as used in the initial evaluation. Model-to-model comparisons will be prepared to clearly identify the differences between ISCST3 and the alternative modeling approach. In addition, we will prepare a series of experimental analyses and assess their usefulness in evaluating modeling for air toxics planning.

3.3 TASK 1: PREPARE A WORK PLAN

This work plan was developed to meet the requirements of this portion of the project. The first step in preparing this work plan was to obtain feedback from stakeholders regarding the types of questions upon which the modeling should be focused. A list of toxics modeling discussion items was developed by STI and sent to LADCO and stakeholders for comment (see Appendix B). The comments provided by stakeholders in response to these discussion items were used to develop the technical approach in this work plan.

3.4 TASK 2: ACQUIRE AND REVIEW DATA

Geophysical, meteorological, air quality, and emissions data are required for this study. Geophysical, meteorological, and emissions data are required to prepare model inputs while air quality data are required for model evaluation. STI will acquire and review all data prior to preparing model inputs.

3.4.1 Geophysical Data

Terrain elevation at each source and receptor are required for ISCST3. CAMx requires land-use data. Digitized terrain elevation and land-use data will be acquired from the U.S. Geological Survey (USGS) if not already available in STI's data library.

3.4.2 Meteorological Data

STI will acquire standard National Weather Service (NWS) meteorological observations (surface and aloft) for Seattle and Detroit for use in the initial air quality simulations. Additional data will be acquired to carry out the full range of simulations. These data include meteorological measurements in the Seattle and Detroit areas from the following sites:

- 10-city pilot monitoring sites,
- routine air monitoring sites within 50 km of the 10-city pilot monitors, and

- NWS and other meteorological monitoring sites within 50 km of the 10-city pilot monitors.

3.4.3 Emissions Data

The Final 1999 National Emissions Inventory Version 3 for Hazardous Air Pollutants will be the principal sources of emissions data for this study. Where alternative toxics inventories (e.g., the Great Lakes Toxic Emission Regional Inventory) are available, their use will be considered. STI will solicit recommendations and comments on emission inventories prior to acquiring and processing them.

3.4.4 Air Quality Data

Toxics data for Seattle and Detroit will be taken from the 10-city pilot study data set acquired and validated in Task 2 of the characterization of air toxics component of this study (Section 2).

3.5 TASK 3: PREPARE MODEL INPUTS

3.5.1 Geophysical

STI will use in-house Geographical Information System (GIS) applications to prepare elevation data inputs for ISCST3 and gridded land-use/surface cover and surface UV albedo codes for CAMx.

3.5.2 Meteorology

The Meteorological Processor for Regulatory Models (MPRM) (U.S. Environmental Protection Agency, 1996) will be used with surface and mixing-height data as input to create ISCST3 input files. Mixing heights will be computed with the mixing-height program provided on the EPA Support Center for Regulatory Modeling web site (http://www.epa.gov/scram001) using surface and upper-air data.

The Diagnostic Wind Model (DWM) (Douglas et al., 1990) and a series of in-house preprocessor programs will be used to prepare gridded, non-steady-state meteorological inputs for CAMx.

3.5.3 Emissions

STI will use the Sparse Matrix Operator Kernel Emissions Modeling System (SMOKE) (Houyoux et al., 2000) to process the toxics emissions data and allocate the data spatially, chemically, and temporally for the following emission types:

- Point: allocated to specific locations
- Non-point: county-level emissions allocated to 1-km grid cells
- On-road mobile: county-level emissions allocated to 1-km grid cells using surrogates for major road segments such as interstates, U.S. highways, and state highways
- Non-road mobile: county-level emissions allocated to 1-km grid cells

SMOKE can produce CAMx emission input files directly. However, for ISCST3, we will prepare a SMOKE postprocessor to reformat the emissions.

3.6 TASK 4: AIR QUALITY MODELING

STI will perform air quality modeling of the 33 urban air toxics listed in Table 2-1 for the cities of Seattle and Detroit. The modeling will be performed in three phases: (1) using routine single-station NWS meteorology as input, (2) using single on-site or nearest-site meteorology as input, and (3) using non-steady-state meteorological fields as inputs. In each case, the air quality model will be applied for a full year corresponding to the periods when toxics measurements were made in each city. A Gaussian plume dispersion model with steady-state meteorology will be used in the first two phases and an Eulerian grid model with non-steady-state meteorology will be used in the last phase. This phased approach allows the model performance to be evaluated incrementally as the level of meteorological definition is increased.

For the first two phases, the ISCST3 model (U.S. Environmental Protection Agency, 1995) will be used as recommended in EPA guidance on urban scale toxics modeling (U.S. Environmental Protection Agency, 1999a).

3.6.1 Chemistry

The ISCST3 model provides concentration estimates due to primary emissions and has a limited capability to consider atmospheric transformations by exponential decay. In addition, some pollutants are formed in the atmosphere due to reactions among other pollutants. Thus, in addition to estimating concentrations due to primary emissions, an estimate of concentrations based on secondary production must be added to the ISCST3 output in order to avoid large underpredictions.

EPA's Ozone Isopleth Plotting for Research (OZIPR) model (Gery and Crouse, 1992) may be used to estimate the secondary transformation of acetaldehyde, formaldehyde, and acrolein. The U.S. Environmental Protection Agency (1999b) describes an approach where secondary toxics production is estimated with the stand-alone OZIPR model that incorporates only nondispersive processes, such as photochemistry. In that approach, the results from this model are then coupled with output from the ISCST3 model, which accounts for dispersion but not for chemical transformation. We will apply the EPA approach in this study.

3.6.2 Background Concentrations

Background air quality includes pollutant concentrations due to natural sources, nearby sources other than those under consideration, and unidentified sources. For typical exposure assessments, background concentrations should be added to the modeled concentrations to provide total ambient air concentrations for estimating exposure. Air quality data from the toxics monitoring and related networks (i.e., PAMS, IMPROVE, CASTNet, NOAA, etc.) will be used to establish background concentrations. Where adequate measurements from these networks are unavailable, background concentrations found in the literature will be used. These will be our primary approaches to establishing background concentrations.

However, in the absence of measured or other reported values, the following approach suggested by EPA may be used. An expanded point source inventory will be obtained for an area surrounding each city. The domain for this expanded point source inventory will extend beyond the domain of the inventory being explicitly modeled in the analysis. An estimate of background concentrations at each receptor within the modeling domain will obtained by multiplying the point source emission rate by a distance dependent factor; sources closer than 50 km are excluded. The modeled background concentration will be based on a summation of concentrations computed from a grid across the modeling domain. The main problem with this approach is that it only provides an estimate of regional contributions to background concentrations and will underestimate concentrations when there are significant continental or global contributions.

3.6.3 ISCST3 Modeling

ISCST3 will be applied for both Seattle and Detroit with receptors defined at toxics monitoring sites and for a grid of receptors spaced 1 km apart. The domains will be designed to encompass all areas within 50 km of the pilot study monitoring sites. ISCST3 will be applied first using standard NWS meteorological data and then using on-site or nearest-site meteorology for successive applications. The actual number of simulations will depend on meteorological data availability and quality.

3.6.4 CAMx Modeling

In a Coordinating Research Council (CRC)-sponsored review of air toxics modeling (Seigneur et al., 2002a), it was recommended that three-dimensional (3-D) grid models be used for toxics at regional and urban scales; at the local scale, plume models should be used to address the near-source impacts of stationary sources, and roadway dispersion models should be used to address the impacts of traffic emissions. Because our objective is to assess the effects of improved meteorological representation on urban-scale analyses, we will apply CAMx initially without toxics chemistry. However, if sufficient budget is available after the initial simulations are completed, a second set of simulations that include chemistry will be performed.

CAMx version 4.03 will be applied with the same 1-km grid spacing used for the ISCST3 receptors, but arranged such that the model's grid cell centers match the ISCST3 receptors.

Bilinear interpolation from the nearest four cell centers will be used to estimate site-specific concentrations.

3.7 TASK 5: EVALUATE MODEL PERFORMANCE

The model results will be compared with the measured data acquired in Task 2 for 18 HAPs using both graphical and statistical methods. The initial evaluation will use the methods from the NATA evaluations, which include scatter plots, ratio box plots, number of sites, median of ratios, percent of sites within a factor of two and within 30%, percent of sites underestimated, and other statistics. The same background concentrations used in NATA will be used initially, but the evaluation will also be performed using measured data if suitable background site data are available. The basic evaluation will be expanded with additional analyses such as spatial concentration plots, plots comparing species, and other focused plots or statistics. The additional analyses will be directed at providing a better understanding of how well the model replicates the measured data and identifying modeled processes that need improvement.

Monitor values below the MDL will be flagged and one-half the MDL will be used for computation of monitor averages. For some pollutant/monitor combinations, many of the monitor readings are below the MDL. For these low-concentration sites, we do not have confidence in the monitored annual averages based on an analysis performed by Battelle. Therefore, only pollutant/monitor combinations with at least 50% of data above the MDL will be included in the evaluations.

3.7.1 Statistical Evaluation

Number of Sites

The number of sites is the number of monitors for each pollutant.

Mean Concentrations

Monthly and annual mean modeled and measured concentrations of each pollutant will be compared.

Bias Statistics

Bias statistics are useful in identifying and quantifying biases in model predictions. Mean absolute bias (ABIAS) provides measure of bias in concentration units. ABIAS is calculated as in Equation 3-1.

$$ABIAS_{threshold} = \frac{1}{N} \sum_{i=1}^{N} Pred^{i}_{x,t} - Obs^{i}_{x,t} \quad \textbf{(3-1)}$$

where N includes all of the model predicted (Pred) and observed (Obs) concentration pairs with observed concentrations above a threshold concentration (e.g., the MDL) for each species, from all stations in a region on a given year. Note the bias is defined as a positive quantity when the model estimate exceeds the observation.

Mean normalized bias (NBIAS) provides a similar metric but normalized to the observed concentrations and given in percent as shown in Equation 3-2.

$$NBIAS_{threshold} = \frac{100}{N} \sum_{i=1}^{N} \frac{(Pred^{i}_{x,t} - Obs^{i}_{x,t})}{Obs^{i}_{x,t}} \quad \textbf{(3-2)}$$

Error Statistics

Error statistics provide estimates of model uncertainty and information complementary to that obtained from bias statistics since unbiased models can still exhibit large errors in their prediction. Mean absolute error (AERROR in concentration units) and mean normalized error (NERROR in percent) calculations are similar to their corresponding bias statistics but employ the absolute value of the model prediction and measured concentration differences. The calculations of AERROR and NERROR are shown in Equations 3-3 and 3-4, respectively.

$$AERROR_{threshold} = \frac{1}{N} \sum_{i=1}^{N} | Pred^{i}_{x,t} - OBS^{i}_{x,t} | \quad \textbf{(3-3)}$$

$$NERROR_{threshold} = \frac{100}{N} \sum_{i=1}^{N} \frac{| Pred^{i}_{x,t} - OBS^{i}_{x,t}|}{Obs^{i}_{x,t}} \quad \textbf{(3-4)}$$

Median of Ratios

The median of ratios is based on the model/monitor ratios for a given pollutant. A median close to one (1) suggests that the model overestimates the monitors about as often as it underestimates the monitors. This statistic is also shown on the ratio box plots.

Percent of Sites Estimated "within a factor of *x*".

This statistic is based on the model/monitor ratios for a given pollutant. We will calculate the percent of sites for a given pollutant which agree within a factor of two, which is

the percent of sites for which the model estimate is somewhere between half and double the monitor average. The percent of sites estimated within 30% (percent of sites for which the model/monitor ratio is between 0.7 and 1.3) will also be calculated.

MAXTOMON

This technique compares the MAXimum model estimate within *r* kilometers of the monitor TO the MONitor average. All model estimates are considered (both estimates at monitor sites as well as the estimates at cell centers) in finding the maximum values. This is an example of a point-to-range technique. This technique is used to test whether biases in the model predictions at monitoring sites are due to location uncertainties or due to systematic model biases.

3.7.2 Graphical Evaluation

Contour Plots

Contour (isopleth) plots of modeled concentrations with monitored values over-plotted will be used to show measured concentrations in the context of modeled concentration gradients. This graphical technique, like the MAXTOMON statistic, allows us to assess whether biases in the model predictions at monitoring sites are due to location uncertainties or due to systematic model biases.

Scatter Plots

Scatter plots are a graphical technique to show the relationship between two variables. Model estimates of annual averages will be plotted against monitor averages. Each ordered pair on the graph is plotted for each monitoring site for that pollutant. Because a much smaller number of monitors is used in this analysis than in a national analysis, daily averages will also be plotted against monitor averages. Comparisons of daily values can also be produced, by month or season, to investigate seasonal variations.

Ratio Box Plots

Ratio box plots show the same data as the scatter plots but in a different fashion. Each box shows the distribution of model-to-monitor ratios. The plots will show the median, 25th, and 75th percentiles of the ratios.

We will use a logarithmic scale for the vertical axis because, if we use a regular arithmetic scale on the vertical axis, then a ratio of 2.0 is twice as far from 1.0 as a ratio of 0.5. However, modelers typically speak of estimates as "within a factor of *x*". An underestimate by a factor of *x* should look just as erroneous as an overestimate by a factor of *x*. A logarithmic scale makes the overestimation and underestimation the same distance from the horizontal line where the ratio is 1.0.

The ratio box plots will be displayed side-by-side, one for each pollutant. This display format will allow us to see easily which toxics are being overestimated and underestimated, and which are being estimated consistently.

3.8 TASK 6: REPORTING AND SYNTHESIS

Throughout the project, we will strive to prepare documents that are usable by readers with a range of backgrounds and knowledge regarding air toxics. The major focus of this task is to synthesize the results from the model evaluation—including inter-model comparisons, site comparisons, and seasonal variations in performance—in a final report, with key findings and recommendations conveyed directly to maximize comprehension. We intend to organize the final results as an executive summary and overview followed by descriptions of modeling and analysis methods, results, and interpretation. The results will also be clearly summarized and explained at a follow-up meeting designated by LADCO in mid-2004.

4. DELIVERABLES AND SCHEDULE

4.1 DELIVERABLES

4.1.1 Data Validation and Analysis

The deliverables for the data validation and analysis tasks include

- Technical memorandum providing data validation guidelines and examples for the states.
- Guidance via telephone and e-mail with CIRA regarding web site development for the EPA.
- Draft and final reports (or other written products selected in consultation with LADCO) characterizing and interpreting the air toxics data, including focused analysis on data from cities with a number of sites (i.e., three or more) with adequate data. These analyses will include meteorological analyses, emission inventory/spatial mapping analyses, and source apportionment analyses, all of which are dependent on the amount, type, and availability of air toxics and supporting data. Analyses will be organized as prescribed analyses for various cities with an overview that describes the approach and ties the common threads together.

4.1.2 Model Evaluation

The deliverables for the model evaluation task include

- Technical memorandum providing the model evaluation results for the ISC3 simulations with recommendations for further modeling and analysis if warranted.
- Draft and final reports summarizing the methods, results, and findings of the model evaluation tasks. These reports will provide recommendations on guidance for future toxics model applications and evaluations. All data, model inputs and outputs, and analyses used in the model evaluation will be provided with the report on CD-ROM.

4.2 SCHEDULE

We estimate the period of performance for this project to be approximately 12 months. We have assumed the following: the toxics database will be provided/acquired with minimal effort and time and in a reasonable format, and we will be able to access the EPA's old Aerometric Information Retrieval System (AIRS) and the new Air Quality System (AQS) data at minimal cost (or obtain necessary data through EPA at no cost). **Table 4-1** summarizes the deliverables and schedule by task. The final report will include an executive summary. For written products, one electronic Microsoft Word copy, one reproducible copy-ready master, and 10 paper copies will be prepared.

Table 4-1. Deliverables and schedule by task.

Task	PI	Deliverable	Due Date
Kick-off conference call	Hilary R. Hafner	Review of scope of work, schedule	August 2003
Finalization of work plan	Hilary R. Hafner	Work plan[a]	December 2003
Progress updates	Hilary R. Hafner	Conference calls	Monthly
Acquisition, review and validation of data	Hilary R. Hafner	Conference calls, DFR	February 2004
Preparation of data displays	Hilary R. Hafner	Conference calls, DFR	February 2004
Model evaluation	Neil J.M. Wheeler	Conference calls Technical Memo DFR	February 2004 April 2004 June 2004
Further analysis and interpretation	Hilary R. Hafner	Conference calls, DFR[b]	June 2004
Presentation of results	Hilary R. Hafner	Presentation at a meeting – likely in Chicago	Summer 2004
Incorporate comments; finalize report	Hilary R. Hafner	FR[b]	Mid-2004

[a] One electronic copy.
[b] One electronic copy and ten paper copies.

Communication is a key component of this work. STI stresses the "big picture" and consensus among analyses (i.e., using several different approaches to meet an analysis objective).

5. MANAGEMENT PLAN

The Principal Investigator (PI) for air quality characterization will be Ms. Hilary Hafner; Mr. Neil Wheeler will be the PI for model evaluation. Ms. Hafner will also be the PI for the overall project management task, including participation in conference calls, supervision of the draft and final reports preparation, and attendance at a meeting to discuss findings. Dr. Paul Roberts and Mr. Frederick Lurmann will serve as the senior technical advisors for the project. Both Ms. Hafner and Mr. Wheeler will be supported by experienced staff. STI will meet the requirements associated with the use of federal funds, and information and data delivered will be in the public domain.

The contract with LADCO for this project is about $250,000, with roughly 75% of funds allocated for characterization analyses and 25% for modeling.

6. REFERENCES

Battelle Memorial Institute and Sonoma Technology, Inc. (2003) Phase II air toxics monitoring data: analyses and network design recommendations. Final technical report prepared for Lake Michigan Air Directors Consortium, Des Plaines, IL by Battelle Memorial Institute, Columbus, OH, and Sonoma Technology, Inc., Petaluma, CA, December.

Bortnick S., Coutant B., Holdren M., Stetzer S., Holdcraft J., House L., Pivetz T., and Main H. (2001) Air toxics monitoring data: Analyses and network design recommendations. Report prepared for Lake Michigan Air Directors Consortium, Des Plaines, IL, by Sonoma Technology, Inc., Petaluma, CA and Battelle Memorial Institute, Columbus, OH, October.

Brown S.G. and Main H.H. (2002) Validation of Massachusetts hazardous air pollutant data. Technical memorandum prepared for the Commonwealth of Massachusetts Dept. of Environmental Protection, Lawrence, MA, by Sonoma Technology, Inc., Petaluma, CA, STI-902131-2218-TM, July 16.

Brown S.G. and Hafner H.R. (2003) Source apportionment of Detroit pilot city air toxics data. *National Workshop on Air Toxics Monitoring, Chicago, IL, May 13-14* Sonoma Technology, Inc., Petaluma, CA.

Brown S.G., Knoderer C.A., and Hafner H.R. (2003) Meteorological effects on air toxics concentrations in Detroit. *National Workshop on Air Toxics Monitoring, Chicago, IL, May 13-14* Sonoma Technology, Inc., Petaluma, CA.

Douglas S.G., Kessler R.C., and Carr E.L. (1990) User's guide for the urban airshed model. Volume III: user's manual for the diagnostic wind model. Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, EPA-450/4-90-007C.

ENVIRON (2003) User's Guide - Comprehensive Air Quality Model with Extensions (CAMx). Version 4.00.

Gery M.W. and Crouse R.R. (1992) User's guide for executing OZIPR. Report prepared by the U.S. Environmental Protection Agency, Research Triangle Park, NC.

Hafner H.R., Funk T.H., Penfold B.M., and Anderson D. (2002) Assessment of the toxics monitoring network in Phoenix, Arizona. Presented at the *Symposium on Air Quality Measurement Methods and Technology, San Francisco, CA, November 13* (STI-901483-2232).

Hafner H.R. (2003a) VOCDat training: validating, reporting and analyzing toxics data. Presented at EPA Region V Offices, Chicago, IL, May 14-15; Research Triangle Park, NC, June 4; Albany, NY, June 10; and EPA Region IX Offices, San Francisco, CA June 17, STI-902600-2365.

Hafner H.R. (2003b) Data validation workshop. Prepared for the Lake Michigan Air Directors Consortium and U.S. Environmental Protection Agency, Region V, Chicago, IL, by Sonoma Technology, Inc., Petaluma, CA, STI-903510-2424, October.

Hafner H.R. and Brown S.G. (2003a) Review of Detroit DRUM impactor data. Technical memorandum prepared for the Lake Michigan Air Directors Consortium, Des Plaines, IL by Sonoma Technology, Inc., Petaluma, CA, STI-902910, February.

Hafner H.R. and Brown S.G. (2003b) Assessment of elemental carbon and its relationship with diesel particulate matter. *National Workshop on Air Toxics Monitoring, Chicago, IL, May 13-14* Sonoma Technology, Inc., Petaluma, CA.

Hafner H.R., Brown S.G., Penfold B.M., and Funk T.H. (2003) Suitability analysis for monitoring diesel particulate matter in the Detroit area. *National Workshop on Air Toxics Monitoring, Chicago, IL, May 13-14* Sonoma Technology, Inc., Petaluma, CA.

Houyoux M., Vukovich J., and Brandmeyer J. (2000) Sparse Matrix Operator Kernel Emissions Modeling System (SMOKE) user manual. Prepared by MCNC-North Carolina Supercomputing Center, Environmental Programs, Research Triangle Park, NC.

Main H.H. and Collins J.F. (1994) Quality assurance, data validation, and data screening for the SCE precipitation chemistry study. Final report prepared for Southern California Edison, Rosemead, CA by Sonoma Technology, Inc., Santa Rosa, CA and Collins Consulting & Computing, Camarillo, CA, STI-92270-1369-FR, April.

Main H.H. (2000) PAMS data validation: VOCDat training. Presented at the *EPA PAMS Data Analysis Workshop, September 10-11* (STI-999434-2008).

Main H.H. and Prouty J.D. (2000) VOCDat user's guide. Prepared for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC by Sonoma Technology, Inc., Petaluma, CA, STI-999434-2015-UG, October.

Main H.H. and Roberts P.T. (2000) PAMS Data Analysis Workbook: Illustrating the use of PAMS data to support ozone control programs. Prepared for the U.S. Environmental Protection Agency, Research Triangle Park, NC, by Sonoma Technology, Inc., Petaluma, CA, STI-900243-1987-FWB, September.

Main H.H. and O'Brien T. (2001) PAMS data validation for the Northeast and Mid-Atlantic states, 1994-1999. Report prepared for NESCAUM, Boston, MA by Sonoma Technology, Inc., Petaluma, CA, STI-900860-2040-FR, June.

Main H.H. and Roberts P.T. (2001) $PM_{2.5}$ Data Analysis Workbook. Draft workbook prepared for the U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, by Sonoma Technology, Inc., Petaluma, CA, STI-900242-1988-DWB, February.

Main H.H. (2002) Phoenix area toxics data 1994-1998: Data validation and analysis summary. Report prepared for the Arizona Department of Environmental Quality, Phoenix, AZ by Sonoma Technology, Inc., Petaluma, CA, STI-901481-2186-FR, May.

Main H.H. and Brown S.G. (2002a) Validation and analysis of Atlanta, Georgia 1999 and 2000 PAMS data. Technical memorandum prepared for the Georgia Department of Natural Resources, Atlanta, GA, by Sonoma Technology, Inc., Petaluma, CA, STI-999434-2177-TM, March.

Main H.H. and Brown S.G. (2002b) PAMS data validation for the northeast states 2000-2001. Report prepared for Northeast States for Coordinated Air Use Management, Boston, MA by Sonoma Technology, Inc., Petaluma, CA, STI-902071-2205-FR, September.

Nizich S. (2001) Data Analysis Workshop. *National Air Toxics Pilot Monitoring Program, Chicago, IL, October 29,* U.S. Environmental Protection Agency Monitoring and Quality Assurance Group, Emissions Monitoring and Analysis Division. Available on the Internet at <http://64.27.125.175/toxics/reports/pdf/overviewepa.pdf>; last accessed November 19, 2003.

Rosenbaum A.S., Ligocki M.P., and Wei Y.H. (1999) Modeling Cumulative Outdoor Concentrations of Hazardous Air Pollutants. Revised Final Report, SYSAPP-99-96/33R2, February.

Ryan P.A., Hafner H.R., and Brown S.G. (2003) Deposition of air pollutants to Casco Bay. Final report prepared for Casco Bay Estuary Project, University of Southern Maine, Portland, ME by Sonoma Technology, Inc., Petaluma, CA, STI-902150-2209-FR2, July.

Ryerson T.B., Trainer M., Nicks D.K., Jr., Neuman A.J., Dissly R.W., Holloway J.S., Jakoubek R.O., Parrish D.D., Frost G.J., Wiedinmyer C., Sueper D.T., Kuster W.C., Goldan P.D., Huebler G., Angevine W.M., Fehsenfeld F.C., Donnelly S.G., Schauffler S., Atlas E., Weinheimer A.J., Flocke F., Wert B.P., Potter W.T., Fried A., Senff C.J., Darby L.S., Banta R.M., and Alvarez R.J. (2001) Contrasting effects of different anthropogenic source types on ozone formation in Houston. Presented at the *Texas 2000 Air Quality Study Workshop in Austin, TX, August 7-10.*

Seigneur C., Pun B., Lohman K., and Wu S. (2002a) Air Toxics Modeling. CP079-02-3.

Seigneur C., Pun B., Lohman K., and Wu S.-Y. (2002b) Air toxics modeling. Report prepared for Coordinating Research Council, Inc., Alpharetta, GA and U.S. Department of Energy's Office of Heavy Vehicle Technologies through the National Renewable Energy Laboratory, Golden, CO by Atmospheric & Environmental Research, Inc., San Ramon, CA, CRC Project Number A-42-1, Document Number CP079-02-3, August. Available on the Internet at <http://www.crcao.com/reports/recentstudies00-02/A-42-1%20Final%20Report.pdf>; last accessed November 19, 2003.

U.S. Environmental Protection Agency (1995) User's guide for the Industrial Source Complex (ISC3) dispersion models, Volumes 1, 2, and 3. Prepared by Pacific Environmental Services, Inc. and U.S. EPA, Research Triangle Park, NC, EPA-454/B-95-003 A, B, and C, September.

U.S. Environmental Protection Agency (1996) Meteorological processor for regulatory models (MPRM) user's guide. Office of Air Planning and Standards, Research Triangle Park, NC, EPA-454/B-96-002, August.

U.S. Environmental Protection Agency (1999a) Air Dispersion Modeling of Toxic Pollutants in Urban Areas: Guidance, Methodology and Example Applications. EPA-454/R-99-021, July.

U.S. Environmental Protection Agency (1999b) A simplified approach for estimating secondary production of hazardous air pollutants (HAPs) using the OZIPR model. Office of Air Planning and Standards, Research Triangle Park, NC, EPA-454/R-99-054.

U.S. Environmental Protection Agency (2000) Air Toxics Monitoring Concept Paper. Peer review draft for the Science Advisory Committee, Air Toxics Monitoring Strategy Subcommittee prepared by Office of Air Quality Planning and Standards, February 29. Available on the Internet at <http://www.epa.gov/ttn/amtic/files/ambient/airtox/cncp-sab.pdf>; last accessed November 29, 2003.

Appendix A
Analysis of Air Toxics Monitoring Data
POLICY QUESTIONS

Sonoma Technology, Inc., under contract to the Lake Michigan Air Directors Consortium, is preparing the workplan for the next phase of air toxics data analysis. In support of this effort, we are trying to understand the policy-oriented questions related to the air toxics data that could potentially be addressed (or, perhaps, have been addressed in previous analyses) with the historical and pilot city data. Roughly grouped by topic area, here are the questions that we have heard are of interest. Are these all of interest? Are the questions accurately stated? Should other questions be added? ***Please provide us with your feedback by September 30, 2003.*** Thanks!

Data Quality

- How good are the historical and pilot city data?
- What are the criteria for determining "good"?
- What is our confidence in the data?
- Which toxics species are adequately represented in the database?
- Are data useful even if they don't meet the QA/QC criteria?
- What are appropriate methods for producing daily, annual, or city-wide averages?
- How should missing data and data below detection levels be treated, and how do different data treatments affect the results?
- Can policy questions be addressed with the available and planned data?
- How do we quantify uncertainty in the data analysis results?

Data Variability

- How do air toxics concentrations vary spatially?
    - Geographic variation: What does a broad national assessment say about air toxics concentrations across the country?
    - City-to-city and intra-urban variation: Based on case studies, what can we say about air toxics concentrations on the urban scale? How do concentrations in urban areas compare with those in nearby rural areas? Are there typical urban and rural profiles (fingerprints)? How does exposure to mobile source toxics vary as a function of distance from roadways? How can concentrations from a small network be extrapolated to other areas? How representative are the existing monitoring sites (i.e., how broadly [spatially] can the annual average concentration at a particular site be applied)?
    - Baseline (background) concentrations: What is meant by "background" levels? How can background levels be estimated? What are reasonable estimates of background levels?
- How do air toxics concentrations vary temporally?
    - What can we say about the variation in air toxics concentrations on a yearly (trends), seasonal (or monthly), day-of-week, and hour-of-day basis?

Control Strategy Effectiveness

- How effective have mobile source controls been in reducing exposure to toxics?
- How effective have maximum achievable control technology (MACT) standards been in reducing exposure to toxics?
- What effect have various community assessment projects had in reducing exposure to air toxics?
- Have emissions controls reduced background concentrations?
- How has implementation of ozone and particulate matter controls benefited air toxics (and vice versa)?
- What is the residual concentration? That is, based on measured data, what is left over from other major toxics mitigation strategies?
- Can air toxics monitoring data be used to help EPA measure GPRA goals (i.e., risk reductions)? [The Government Performance and Results Act of 1993 (GPRA) requires that federally funded agencies develop and implement an accountability system based on performance measurement, including setting goals and objectives and measuring progress toward achieving them.]

Appendix B
ANALYSIS OF AIR TOXICS MONITORING DATA
MODEL EVALUATION COMPONENT DISCUSSION ITEMS

Overarching Requirements for the Project

Use/build from existing analyses

Use both historical national data set and pilot city data set

Let the data tell the story

Clearly communicate policy relevant scientific information

The Role of Models in Toxics Analysis

Air quality models are mathematical tools that simulate atmospheric processes, including transport, diffusion, transformation, and removal. Air quality models have been used for decades to ***assess the potential impact of emission sources*** on ambient concentrations of criteria and toxic air pollutants. In the past decade, air quality models have also been used as ***planning tools*** for assessing the effectiveness of emission controls in reducing criteria pollutant concentrations. However, until recently, air quality models have not been used as planning tools for toxics, due to the lack of measurements with which to evaluate the models. There is a clear need to assess the usefulness of these models in air quality planning and to improve both modeling and evaluation methods. One of the objectives of the national toxics monitoring program is to provide data to support and evaluate dispersion and deposition models.

Potential Questions to Answer through Modeling

1. How good is the modeling approach that was used in the NATA?

Background: The modeling with ASPEN performed under the NATA did not show good agreement with observations. It has been suggested that the model, the emission inventory, and/or the toxics measurements may be the reason(s) for poor agreement between measurements and model estimates. There is a need to assess if the NATA modeling results are useful in light of better measurements and emissions estimates.

Approach: Perform ASPEN modeling following the NATA modeling and evaluation protocol using updated emissions and toxics data from one or more pilot cities.

Deliverables: An assessment of the usefulness of ASPEN for impact assessment and planning (Is it good enough for assessment? Is it good enough for planning?).

Issues: The pilot cities with the best temporal and spatial coverage of toxics data (Seattle and Detroit) may have more complex meteorology than other cities because of land-water interfaces. The adequacy of background concentrations still needs to be addressed.

2. What is the impact of Emission Inventory Quality on predicted concentrations?

Background: It is known that there are significant uncertainties in toxics emission inventories. As Joe Touma indicated in our conference call with EPA, work is underway to prepare improved inventories with greater spatial, temporal, and chemical resolution; and to model toxics with those improved inventories and more complex models (i.e., ISCST3). As Joe also noted, these emission inventory improvements are expensive and would not be within the budget of this project.

Issues: Even if emission inventories are perfect, model formulation and meteorological inputs may be inadequate.

3. In the absence of adequate background measurements how can background values for toxics be defined for use in near-source modeling?

Background: Toxics monitoring has historically focused on areas where there is likely to be higher levels of human exposure. This focus has not provided adequate measurements to establish background concentrations of toxics for use in modeling assessments. Establishing boundary conditions is complicated because some important toxics are VOCs (e.g., benzene, formaldehyde, 1,3-butadiene, and acetaldehyde) and undergo chemical transformation in the atmosphere. Further complications exist because formaldehyde and acetaldehyde can be formed from every VOC in the atmosphere. Major contributors are toluene, xylenes, auto exhaust, and biogenic hydrocarbons. It is estimated that more than 80% of these aldehydes may be due to atmospheric formation, not emissions (Luecken, 2002). Therefore, there is no good way to predict aldehyde concentrations with a dispersion model alone.

Approach: Perform regional photochemical grid modeling, using Eta Data Assimilation System (EDAS) meteorology and national emission inventories as input to derive toxics background concentrations for one or more cities and compare to measurements. Integrate findings with data analysis investigations.

Deliverables: Guidance on establishing background concentrations for toxics modeling and an evaluation of the methodology investigated.

Issues: Not all of the important chemical reactions are represented in current chemical mechanisms used in photochemical grid models.

4. What is the impact of complex meteorology and terrain on predicted concentrations?

Background: While ISCST3 is superior to the ISCLT-like formulation of ASPEN, its meteorological inputs are spatially invariant and cannot be used successfully to model source-

receptor relationships in areas where there is a significant variation in winds. This was one of the reasons CARB commissioned the development of CALMET-CALPUFF in the late 1980s to support their toxics program.

Approach: Develop non-steady-state meteorological fields for one or more pilot cities using CALMET, model toxics using CALPUFF or CALGRID, evaluate model performance, and compare performance to ASPEN modeling for the same cities. Integrate guidance developed under Question 3.

Deliverables: An assessment of incremental improvement in model performance due to more complete representations of meteorology.

Issues: Available meteorological data may be inadequate to resolve spatial variations of winds in the cities modeled.

5. Can current models predict the range, variability, and gradients of measured toxics?

Background: To date, assessments of toxics modeling has focused on the ability of models to replicate annual concentrations at specific monitoring sites. However, for use in planning it may be more important that models replicate the range, variability, and gradients of toxics.

Approach: Re-evaluate the modeling performed under Questions 1 and/or 4 with an emphasis on local and regional toxics concentration variability. Compare with traditional model evaluation metrics.

Deliverables: Expanded guidance on evaluating toxics modeling.

References

Luecken, D. (2002) Atmospheric Formation and Decay of Air Toxics – Implications for Exposure Assessments. Presented at the Air Toxics Exposure Assessment Workshop, June 25-27, 2002, San Francisco, CA.