Sonoma Technology, Inc.

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

ANALYTICAL TECHNIQUES FOR TECHNICAL ASSESSMENTS OF AMBIENT AIR MONITORING NETWORKS

GUIDANCE DOCUMENT
STI-905104.02-2805-GD

By:
Sean M. Raffuse
Dana C. Sullivan
Michael C. McCarthy
Bryan M. Penfold
Hilary R. Hafner
Sonoma Technology, Inc.
Petaluma, CA

Prepared for:
U.S. Environmental Protection Agency
Research Triangle Park, NC

September 2005

Sonoma Technology, Inc.

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

ANALYTICAL TECHNIQUES FOR TECHNICAL ASSESSMENTS OF AMBIENT AIR MONITORING NETWORKS

GUIDANCE DOCUMENT
STI-905104.02-2805-GD

By:
Sean M. Raffuse
Dana C. Sullivan
Michael C. McCarthy
Bryan M. Penfold
Hilary R. Hafner
Sonoma Technology, Inc.
1360 Redwood Way, Suite C
Petaluma, CA 94954-1169

Prepared for:
Office of Air Quality Planning and Standards
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

September 30, 2005

ACKNOWLEDGMENTS

The authors thank Kevin Cavender, the work assignment manager, for his assistance with this project. We also acknowledge those involved with the National Analysis (Schmidt, 2001) and the preliminary regional assessments, upon which this document draws.

FOREWORD

This document represents guidance for the assessment of technical aspects of air pollutant monitoring networks. It is designed to be flexible and expandable with additional types of analyses and examples as techniques are improved, enhanced, and more broadly applied. Its intended audience includes regional, state, and local air quality planning agencies. Depending on their unique situations, users of this guidance may select one or more analyses, or they may creatively modify one of the recommended analyses to facilitate a monitoring network assessment.

The contents of this document are summarized briefly in the following paragraphs:

- Section 1 summarizes the context of network assessments in general and this specific document, including background and key issues. In addition, Section 1 provides an overview of the procedure for network assessment as described in the draft National Ambient Air Monitoring Strategy (NAAMS) (U.S. Environmental Protection Agency, December 2005).
- Section 2 expands on the procedure for network assessment described in the NAAMS. It introduces consideration of the purposes of a monitoring network—i.e., a network's mission. The purposes provide a basis for performing a network assessment. They are the benchmarks against which the strengths and weaknesses of the network are measured. Section 2 continues with specific details for technical approaches to network assessments, including three general categories of analyses: site-by-site, bottom-up, and network optimization.
- Section 3 expands on the technical approaches introduced in Section 2. It includes a selection of two-page illustrations of analyses for network assessments.
- Section 4 concludes this guidance document with a summary and recommendations for further development of network assessment guidance with an emphasis on expected results and resource requirements. More detailed descriptions of the techniques, and more examples, could be added to future versions of this document as techniques are refined and more broadly applied.
- Section 5 lists the references cited in this guidance document.
- Appendix A discusses project-level example applications of the technical approaches discussed in Sections 2 and 3.
- Appendix B provides examples of projects, completed by Sonoma Technology, Inc., in which three of the more complex analysis techniques discussed in Sections 2 and 3 were used.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The U.S. Environmental Protection Agency (EPA) prepared a draft National Ambient Air Monitoring Strategy (NAAMS) in December 2005 (U.S. Environmental Protection Agency, 2004). The purpose of the NAAMS is to optimize U.S. air monitoring networks to achieve, with limited resources, the best possible scientific value and protection of public and environmental health and welfare. An important element of NAAMS is a plan for periodic network assessments at the national, regional, and local scales. A network assessment includes (1) re-evaluation of the objectives and budget for air monitoring, (2) evaluation of a network's effectiveness and efficiency relative to its objectives and costs, and (3) development of recommendations for network reconfigurations and improvements. Initial network assessments for the NAAMS were led by EPA and its 10 regional offices in 2001 through 2004 (U.S. Environmental Protection Agency, 2003b). This initial assessment, as well as peer-reviews of the NAAMS by subcommittees of the EPA Clean Air Scientific Advisory Committee (Hopke, 2003; Henderson, 2005), produced the recommendation that guidance for regional-scale network assessments be established.

The draft of the NAAMS, and documentation of the initial national- and regional-scale network assessments provide a valuable context and a summary of the key technical issues for network assessment guidelines. This document builds on the lessons learned in the NAAMS and focuses on providing guidance on analytical techniques that can be used for regional-scale assessments.

1.1 BACKGROUND AND KEY ISSUES

Ambient air monitoring objectives have shifted over time—a situation which has induced air quality agencies to re-evaluate and reconfigure monitoring networks. A variety of factors contribute to these shifting monitoring objectives:

- Air quality has changed—for the better in most geographic areas—since the adoption of the federal Clean Air Act and National Ambient Air Quality Standards (NAAQS). For example, the problem of high ambient concentrations of lead has largely been solved.
- Populations and behaviors have changed. For example, the U.S. population has (on average) grown, aged, and shifted toward urban and suburban areas over the past four decades. In addition, rates of vehicle ownership and annual miles driven have grown.
- New air quality objectives have been established, including rules to reduce air toxics, fine particulate matter ($PM_{2.5}$)[1], and regional haze.
- The understanding of air quality issues and the capability to monitor air quality have both improved. Together, the enhanced understanding and capabilities can be used to design more effective air monitoring networks.

As a result of these changes, air monitoring networks may have unnecessary or redundant monitors or ineffective and inefficient monitoring locations for some pollutants, while other

[1] Particulate matter of less than 2.5 microns aerodynamic diameter.

regions or pollutants suffer from a lack of monitors. Air monitoring agencies should, therefore, refocus monitoring resources on pollutants that are new or persistent challenges, such as $PM_{2.5}$, air toxics, and ground-level ozone, and should deemphasize pollutants that are steadily becoming less problematic and better understood, such as lead and carbon monoxide (CO). In addition, monitoring agencies need to adjust networks to protect today's population and environment, while maintaining the ability to understand long-term historical air quality trends. Moreover, monitoring networks can take advantage of the benefits of new air monitoring technologies and improved scientific understanding of air quality issues. Existing monitoring networks should be designed to address multiple, interrelated air quality issues and to better operate in conjunction with other types of air quality assessments (e.g., photochemical modeling, emission inventory assessments). Reconfiguring air monitoring networks can enhance their value to stakeholders, scientists, and the general public.

1.2 OVERVIEW OF MONITORING NETWORK TECHNICAL ASSESSMENTS

Analytical techniques to assess the technical aspects of monitoring networks fit within the overall framework of regional network assessments discussed in the draft NAAMS (U.S. Environmental Protection Agency, 2005). The NAAMS briefly describes the stepwise procedure for network assessments shown in **Table 1-1**. This document focuses on Steps 3 and 4: statistical analyses and objective situational analyses.

In some cases, network assessments can be handled simply by answering one or more straightforward questions. In others, detailed analytical techniques, such as those discussed in Section 2.2 and Section 3, are necessary. A thorough technical assessment will help inform decisions about reconfiguring a network. These decisions might include eliminating redundant monitors, reducing or expanding the monitoring season, moving monitors to better locations, switching a site to different technology (e.g., finer temporal resolution), adding monitors to the network, or switching a site to a different pollutant. In practice, a combination of several types of analyses might provide the most useful information.

Table 1-1. Descriptions and examples of steps involved in performing network assessments.

Step	Description	Examples
1	Prepare or update a regional description, discussing important features that should be considered for network design	Topography, climate, population, demographic trends, major emissions sources, and current air quality conditions
2	Prepare or update a network history that explains the development of the air monitoring network over time and the motivations for network alterations, such as shifting needs or resources.	Historical network specifications (e.g., number and locations of monitors by pollutant and by year in graphical or tabular format); history of individual monitoring sites
3	Perform statistical analyses of available monitoring data. These analyses can be used to identify potential redundancies or to determine the adequacy of existing monitoring sites.	Site correlations, comparisons to the NAAQS, trend analysis, spatial analysis, and factor analysis
4	Perform situational analyses, which may be objective or subjective. These analyses consider the network and individual sites in more detail, taking into account research, policy, and resource needs.	Risk of future NAAQS exceedances, demographic shifts, requirements of existing state implementation plans (SIP) or maintenance plans, density or sparseness of existing networks, scientific research or public health needs, and other circumstances (such as political factors)
5	Suggest changes to the monitoring network on the basis of statistical and situational analyses and specifically targeted to the prioritized objectives and budget of the air monitoring program.	Reduction of number of sites for a selected pollutant, enhanced leveraging with other networks, and addition of new measurements at sites to enhance usefulness of data
6	Acquire the input of state and local agencies or stakeholders and revise recommendations as appropriate	

2. APPROACH TO MONITORING NETWORK TECHNICAL ASSESSMENTS

This section provides guidance to the user for identifying monitoring needs and introduces network assessment analyses.

2.1 IDENTIFY MONITORING NEEDS

Before beginning a network assessment, the purposes of the network must be revisited and prioritized. Networks are likely to be used to meet a variety of purposes, such as monitoring compliance with the NAAQS, assessment of population exposure to pollutants, assessment of pollutant transport, monitoring of specific emissions sources, monitoring of background conditions, and possibly others. These purposes may be prioritized as primary or secondary and individual monitors within a network may serve different purposes. Each analytical technique selected to support a network assessment must be chosen in view of the purposes of the overall network and its individual monitoring sites. In addition, the resources invested in each analysis should be proportional to the priority of the purposes that are being evaluated. **Table 2-1** briefly lists some typical purposes for monitoring networks, although this list is neither comprehensive nor universally applicable to all pollutants.

Network assessments quantifiably measure the successes and shortcomings of monitoring networks' capabilities to meet their monitoring purposes. Therefore, clearly defined monitoring purposes are the basis for the technical assessment of a monitoring network. Once the purposes are defined, appropriate statistical or situational analyses may be considered and selected to evaluate each.

Table 2-1. Typical purposes for ambient air monitoring networks.

Page 1 of 3

Purpose	Examples	Comments
Establish regulatory compliance	Meet national requirements	Monitors may be sited to address NAAQS compliance or may be mandated by prior regulations or SIP provisions.
	Meet state and local regulations	States, or local air districts, may have air quality regulations that are more stringent than federal requirements.
Develop scientific understanding of air quality by supporting other types of assessments or analyses	Air quality model evaluation	Monitors near modeling domain boundaries are useful for defining boundary conditions. Monitors throughout a domain assist model application and evaluation.
	Emission reduction evaluation or emission inventory evaluation	Urban core and maximum emission area monitors can be helpful for evaluating inventories and tracking emissions.

Table 2-1. Typical purposes for ambient air monitoring networks.

Purpose	Examples	Comments
Develop scientific understanding of air quality by supporting other types of assessments or analyses (continued)	Source apportionment	Monitors collecting data on many species (e.g., speciated $PM_{2.5}$) and at fairly high time resolution (1-in-3-day or better) are useful for source apportionment analyses.
	Temporal variability	Sub-daily (e.g., 1-hr, 3-hr) data can be used to track diurnal patterns.
Understand historical trends in air quality	Trend tracking	Monitors with long histories are valuable for understanding and tracking long-term trends.
	Historical consistency	Monitoring sites whose sampling methods have not been changed help maintain consistency for annual comparisons.
Characterize specific geographic locations or emissions sources	Monitor the air quality impacts of an emissions source	Monitors located close to specific source hot spots are useful for tracking emissions from a particular source and developing emission reduction strategies or tracking changes due to controls.
	Monitor the area of maximum precursor emissions	For secondary pollutants such as ozone, monitors located in areas of maximum precursor emissions are useful for modeling and control strategy design.
	Monitor the area of maximum pollutant concentration	Monitors located downwind of maximum emissions.
	Monitor the background concentration	Properly sited background monitors routinely measure the lowest expected values in the region. These monitors are used to asses regional vs. local contributions.
	Monitor surrogate pollutants	Some measurements are useful as surrogates for other pollutants that are not widely monitored. For example, CO monitors can be used as surrogates for wood smoke (Park et al., 2005).

Table 2-1. Typical purposes for ambient air monitoring networks.

Page 3 of 3

Purpose	Examples	Comments
Track the spatial distribution of air pollutants	Transport/border characterization	Sites located near political boundaries or between urban or industrial areas are useful for characterizing transport of pollutants between jurisdictions.
	Interpolation and understanding pollutant gradients	High monitor density improves interpolation maps such as those used in AIRNow (U.S. Environmental Protection Agency, 2003a). Monitors near the urban boundary are particularly useful for constraining the interpolation of high concentrations.
	Forecasting assistance	Upwind monitors are useful for air quality forecasting. For forecasting ozone, NO_x measurements are helpful. For $PM_{2.5}$, continuous monitors are very valuable.
Evaluate population exposures to air pollutants	Environmental justice	Monitors located in areas that have large low income and/or minority populations may be of particular value for assessing environmental justice issues.
	Sensitive groups	Monitors located where people live, work, and play are important for addressing exposure and protecting public health.

2.2 METHODS FOR TECHNICAL ASSESSMENT

2.2.1 Overview

In this document, techniques for assessing technical qualities of monitoring networks are grouped into three broad categories: site-by-site, bottom-up, and network optimization. Site-by-site comparisons rank individual monitors according to specific monitoring purposes; bottom-up analyses examine data other than ambient concentrations to assess optimal placement of monitors to meet monitoring purposes; and network optimization analyses evaluate proposed network design scenarios. Within these broad categories, specific techniques are rated by their complexity on the following scale.

- * Minimal special skills needed; quick
- ** May require common tools, readily available data, and/or basic analysis skills; quick
- *** Requires analysis skills; moderate investment of time
- **** Significant analytical skills, specialized tools; time-intensive or iterative

2.2.2 Site-By-Site Analyses

Site-by-site analyses are those that assign a ranking to individual monitors based on a particular metric. These analyses are good for assessing which monitors might be candidates for modification or removal. Site-by-site analyses do not reveal the most optimized network or how good a network is as a whole. In general, the metrics at each monitor are independent of the other monitors in the network.

Several steps are involved in site-by-site analysis:

1. Determine which monitoring purposes are most important
2. Assess the history of the monitor (including original purposes)
3. Select a list of site-by-site analysis metrics based on purposes and available resources
4. Weight metrics based on importance of purpose
5. Score monitors for each metric
6. Sum scores and rank monitors
7. Examine lowest ranking monitors for possible resource reallocation

The low-ranking monitors should be examined carefully on a case-by-case basis. There may be regulatory or political reasons to retain a specific monitor. Also, the site could be made potentially more useful by monitoring a different pollutant or using a different technology. **Table 2-2** lists specific site-by-site analysis techniques, which are summarized in greater detail in Section 3.

Table 2-2. Site-by-site analysis techniques.

Page 1 of 2

Technique	Complexity	Objectives Assessed (See Table 2-1)	Summary Page
Number of other parameters monitored at the site	*	Overall site value Model evaluation Source apportionment	3-3
Trends impact	* to **	Trend tracking Historical consistency Emission reduction evaluation	3-5
Measured concentrations	**	Maximum concentration location Model evaluation Regulatory compliance Population exposure	3-7
Deviation from NAAQS	**	Regulatory compliance Forecasting assistance	3-9
Area served	**	Spatial coverage Interpolation Background concentration	3-11

Table 2-2. Site-by-site analysis techniques.

Page 2 of 2

Technique	Complexity	Objectives Assessed (See Table 2-1)	Summary Page
Monitor-to-monitor correlation	** to ***	Model evaluation Spatial coverage Interpolation	3-13
Population served	***	Population exposure Environmental justice	3-19
Principal component analysis	***	Background concentration Forecasting assistance	3-21
Removal bias	***	Regulatory compliance Model evaluation Spatial coverage Background concentration Interpolation	3-23

2.2.3 Bottom-Up Analyses

Bottom-up methods examine the phenomena that are thought to cause high pollutant concentrations and/or population exposure, such as emissions, meteorology, and population density. For example, emission inventory data can be used to determine the areas of maximum expected concentrations of pollutants directly emitted (i.e., primary emissions). Emission inventory data are less useful to understand pollutants formed in the atmosphere (i.e., secondarily formed pollutants). Multiple data sets can be combined using spatial analysis techniques to determine optimum site locations for various objectives. Those optimum locations can then be compared to the current network. In general, bottom-up analyses indicate where monitors are best located based on specific objectives and expected pollutant behavior. However, bottom-up techniques rely on a thorough understanding of the phenomena that cause air quality problems. The most sophisticated bottom-up analysis techniques are complex and require significant resources (time, data, tools, and analytical skill). **Table 2-3** lists the specific bottom-up analysis techniques detailed in Section 3. Site-by-site and bottom-up analyses are complementary.

Table 2-3. Bottom-up analysis techniques.

Technique	Complexity	Objectives Assessed (Table 2-1)	Summary Page
Emission Inventory	** to ****	Emission reduction evaluation Maximum precursor location	3-15
Population density	**	Population exposure Environmental justice	Not included
Population change	***	Population exposure Environmental justice Maximum precursor location	3-17
Suitability modeling	****	Population exposure Environmental justice Source-oriented Model evaluation Maximum concentration location Background concentration Transport/border characterization	3-25
Photochemical modeling	****	Maximum concentration location Source-oriented Transport/border characterization Population exposure Background concentration	Not included

2.2.4 Network Optimization Analyses

Network optimization techniques are a holistic approach to examining an air monitoring network. These techniques typically assign scores to different network scenarios; alternative network designs can be compared with the current (base-case) design.

An example of a network optimization analysis is the EPA Region 3 ozone network reassessment (Cimorelli et al., 2003). Region 3 utilized an iterative 10-step process:

1. Select the set of scenarios (i.e., different hypothetical network designs) to be ranked
2. Define decision criteria for scoring each network design
3. Gather the data necessary to calculate scores for the decision criteria
4. Index decision criteria to a common scale
5. Weight the criteria based on relative importance
6. Produce initial results (ranking of scenarios)
7. Iterate – adjust scenarios, decision criteria, and criteria weighting as new information and understanding are developed
8. Obtain feedback from stakeholder deliberation
9. Finalize network optimization scenario results
10. Recommend changes

The formal analytical process used by Region 3 is called Multi-Criteria Integrated Resource Assessment (MIRA) (Cimorelli et al., 2003; Stahl et al., 2002). Forty metrics were used as decision criteria in the analysis. These metrics were arranged hierarchically with four top-level criteria: air quality, personnel impact, costs, and trends impact. For assessing the air quality criteria, Region 3 developed a base-case ozone concentration grid using photochemical modeling results.

Many of the metrics used by Region 3 in their assessment are similar to the analyses described as "site-by-site" analyses in this document. When different network scenarios are considered, the individual monitor scores for a particular analysis can be summed to provide a total score for the entire network. The total score can be compared to other network designs. **Table 2-4** lists some techniques for network optimization. Further details are provided in Section 3.

Table 2-4. Network optimization analysis techniques.

Technique	Complexity	Objectives Assessed (Table 2-1)	Summary Page
Monitor-to-monitor correlation	** to ***	Model evaluation Spatial coverage Interpolation	3-13
Principal Component Analysis	***	Background concentration Forecasting assistance	3-21
Removal bias	***	Regulatory compliance Model evaluation Spatial coverage Background concentration Interpolation	3-23

3. METHOD SUMMARY SHEETS

The following pages are organized into summary sheets for individual analysis techniques. The sheets are designed to provide the vital statistics for the techniques at a glance and to help the analyst narrow down the list of possible analyses to perform based on their available resources and objectives. These sheets cover a range of analysis techniques that can be applied to network assessment; these summaries can be expanded and others can be prepared as examples become available.

Figure 3-1 shows an example front page of a summary sheet. The front page of a summary sheet contains basic information about the type of analysis, the objectives that can be assessed, and the complexity and resources required. It also lists some advantages and disadvantages of the analysis and provides a list of other analyses that can provide similar information but may be more or less complex. The back page provides more detail about the technique, including an example, interpretation, and references for more information.

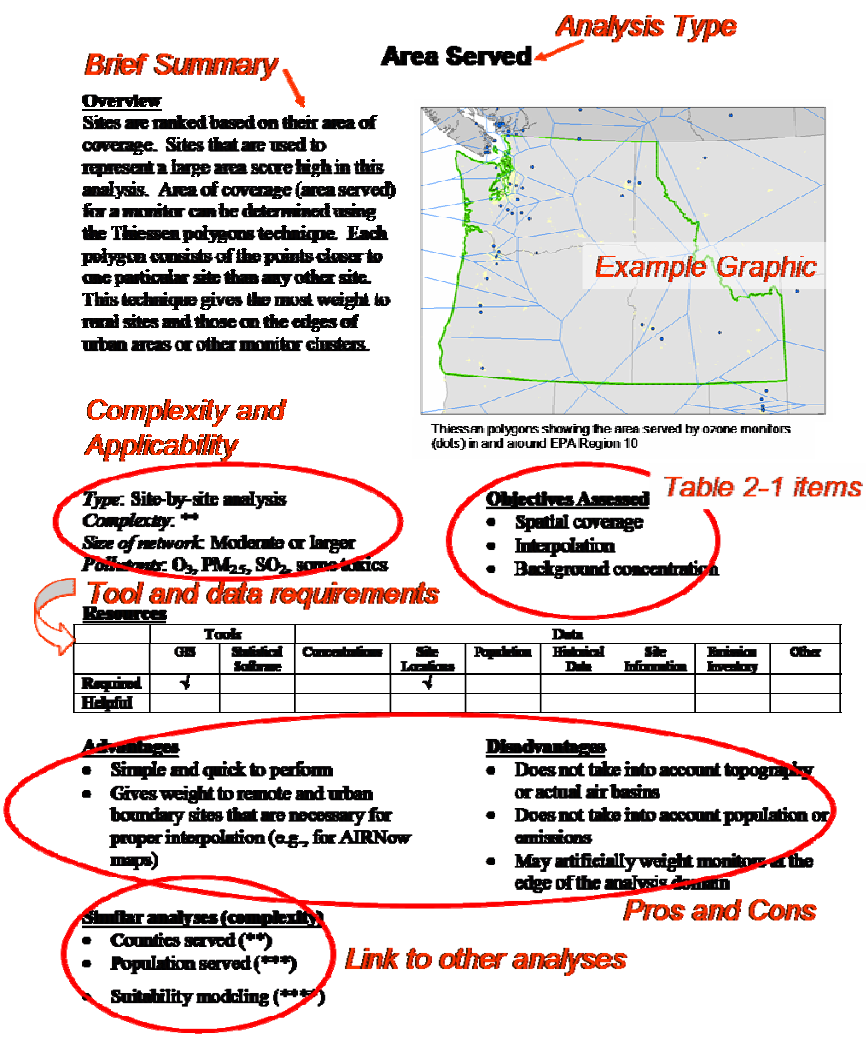
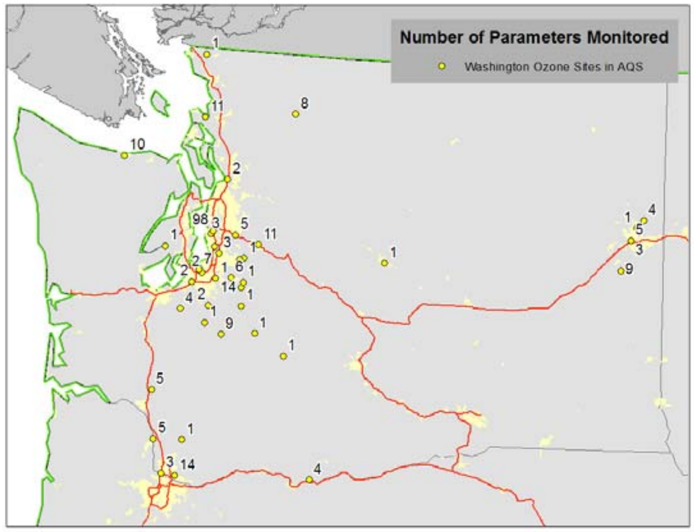


Figure 3-1. Summary sheet front page example.

Number of Other Parameters Monitored

Overview

Monitors that are collocated with other measurements at a particular air quality site are likely more valuable than sites that measure fewer parameters, particularly for source apportionment and other air quality studies. In addition, the operating costs can be leveraged among several instruments at these sites. Sites are ranked by the number of parameters (or instruments) that are collected at the particular site.



Count of additional parameters measured at Washington ozone sites within AQS

Type: Site-by-site analysis
Complexity: *
Size of network: any
Pollutants: any

Objectives Assessed

- Overall site value
- Model evaluation
- Source apportionment

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required				✓			✓		
Helpful									

Advantages

- Simple to perform (given data)
- Good first step in understanding monitor sites

Disadvantages

- Method does not “weight” the measurements (some pollutant measurements may be more useful than others)
- Up-to-date information on the pollutants measured at particular sites can be difficult to acquire

Similar Analyses (Complexity)

None

Number of Other Parameters Monitored

Analysis Goals

This analysis is performed by simply counting the number of other parameters that are measured at the physical site. Sites with many parameters measured are ranked highest. The metric addresses two aspects of monitor value. First, collocated measurements of several pollutants are valuable for many air quality analyses, such as source apportionment, model evaluation, and emission inventory reconciliation. Second, having a single site with multiple measurements is more cost-effective to operate than having monitors scattered at several sites. Other cost-based metrics were included in the Region 3 2003 network assessment.

Example

This example in and around the Seattle, Washington, area was created in ESRI ArcGIS 9.1, using the following steps:

1. Download monitor information from the Air Quality System (AQS) database.
2. Use the monitor coordinate information to determine which monitoring sites are within the study domain.
3. Sum the monitoring (measurements) parameters for each monitor location and determine the best locations to utilize in future air quality studies.

Interpretation

The table at right is an extract of the analysis example for Seattle. The monitor locations are ranked by the number of parameters measured. As shown in the table, three monitors are located within the project study domain and measure numerous parameters. The site measuring 98 parameters is the most valuable for scientific analyses, such as emission inventory reconciliation and source apportionment.

AIRS Code	Number of Parameters Measured	Study Domain
530330080	98	✓
530110011	14	
530330023	14	✓
530330017	11	✓
530570018	11	
530090012	10	
530630001	9	

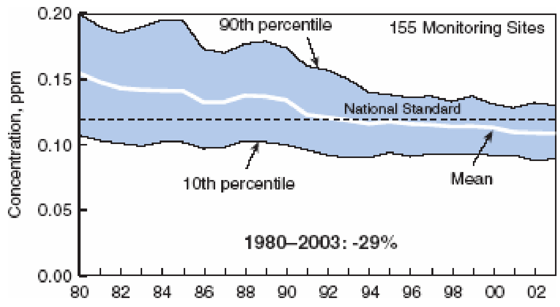
References

Cimorelli A.J., Chow A.H., Stahl C.H., Lohman D., Ammentorp E., Knapp R., and Erdman T. (2003) Region III ozone network reassessment. Presented at the *Air Monitoring & Quality Assurance Workshop, Atlanta, GA, September 9-11* by the U.S. Environmental Protection Agency, Region 3, Philadelphia, PA. Available on the Internet at http://www.epa.gov/ttn/amtic/files/ambient/pm25/workshop/atlanta/r3netas.pdf last accessed September 9, 2005.

Trends Impact

Overview

Monitors that have a long historical record are valuable for tracking trends. In this analysis, sites are ranked based on the duration of the continuous measurement record. The analysis can be as simple as ranking the available monitors based on the length of the continuous sampling record. This technique places the most importance on sites with the longest continuous trend record.



National ozone trends from EPA ozone trend report, 2003.

Type: Site-by-site analysis
Complexity: * to **
Size of network: any
Pollutants: any

Objectives Assessed

- Trend tracking
- Historical consistency
- Emission reduction evaluation

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required						✓			
Helpful			✓						

Advantages

- Simple analysis, requiring few statistical tools
- Useful for identifying long-term trend sites
- A good first look at monitor history

Disadvantages

- Length of continuous record does not ensure that data are of good quality throughout the time period
- Magnitude or direction of past trends are not necessarily good predictors of future trends
- Does not take into account changes in population, emissions, or meteorology

Similar Analyses (Complexity)

- Number of other parameters monitored (*)
- Measured concentrations (**)
- Deviation from NAAQS (**)

Trends Impact

Analysis Goals

Determining the trends impact of a monitor can be done simply. One approach is to rank sites based on their length of continuous sampling. Sites with the longest term of operation would score higher than those with shorter records, since they would be more useful for long-term trend analysis. Additional factors that could be used to adjust the simple ranking scale include (1) the magnitude and direction of trends observed to date at the site, (2) the suitability of a site's location for monitoring trends after a significant event (e.g., enactment of a specific control measure), or (3) proximity of another monitor that could be used to continue the trend record. A site may be weighted as less important if changes in sampling and analysis methodology lead to a discontinuous record. Weighing these factors would require consideration of the overall goals of the monitoring network and the importance of the historical record.

Example

This table shows the number of annual averages available for tetrachloroethylene at toxics trends sites from 1990 to 2003. For this analysis, sites with the longest record would be rated higher than those with shorter records.

City, State	AQS SiteID	Years
Stockton, CA	06-077-1002	13
Baltimore, MD	24-510-0040	12
Los Angeles, CA	06-037-1002	11
San Francisco, CA	06-001-1001	10
Fresno, CA	06-019-0008	10
Baltimore, MD	24-005-3001	10
Los Angeles, CA	06-037-1103	9
Los Angeles, CA	06-037-4002	9
San Diego, CA	06-073-0003	9
San Francisco, CA	06-075-0005	9
San Jose, CA	06-085-0004	9
Baltimore, MD	24-510-0006	9
Sacramento, CA	06-061-0006	8
San Diego, CA	06-073-0001	8
Oxnard, CA	06-111-2002	8
Chicago, IL-IN-WI	18-089-2008	8
Baltimore, MD	24-510-0035	8

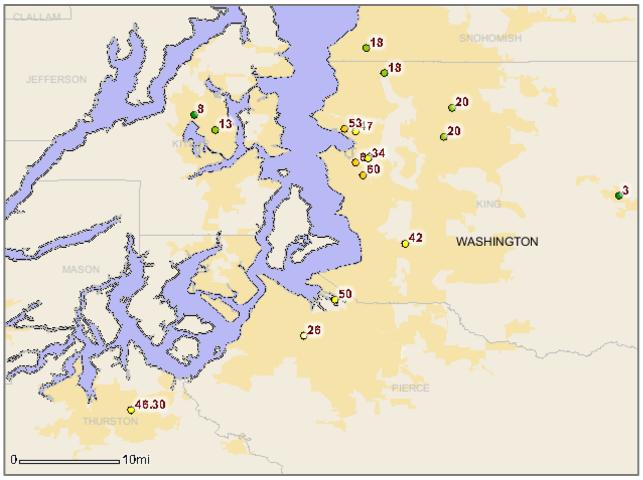
References

Cimorelli A.J., Chow A.H., Stahl C.H., Lohman D., Ammentorp E., Knapp R., and Erdman T. (2003) Region III ozone network reassessment. Presented at the *Air Monitoring & Quality Assurance Workshop, Atlanta, GA, September 9-11* by the U.S. Environmental Protection Agency, Region 3, Philadelphia, PA. Available on the Internet at http://www.epa.gov/ttn/amtic/files/ambient/pm25/workshop/atlanta/r3netas.pdf last accessed September 9, 2005.

Measured Concentrations

Overview

Individual sites are ranked based on the concentration of pollutants they measure. Monitors that measure high concentrations or design values are ranked higher than monitors that measure low concentrations. Results can be used to determine which monitors are less useful in meeting the selected objective.



1-hour $PM_{2.5}$ concentrations in the Seattle area (ug/m^3)

Type: Site-by-site analysis
Complexity: **
Size of network: any
Pollutants: any above detection limits

Objectives Assessed

- Maximum concentration location
- Model evaluation
- Regulatory compliance
- Population exposure

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required			✓						
Helpful		✓					✓		

Advantages

- Identifies key sites from a regulatory perspective based on maximum concentrations.

Disadvantages

- Does not account for monitor-siting problems; monitors may not be measuring maximum concentrations if not properly placed.
- Only focuses on high concentrations; low-concentration monitors may be useful for representing rural locations or background concentrations.

Similar Analyses (Complexity)

Deviation from NAAQS (**)
Emission inventory (** to ****)

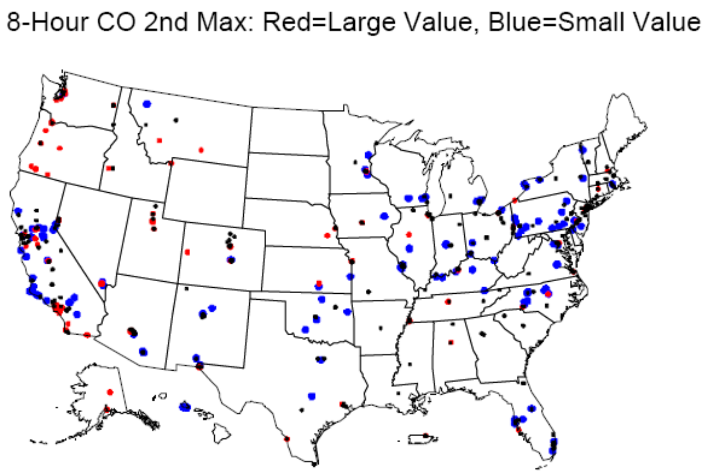
Measured concentrations

Analysis Goals

Sites that measure high concentrations are important for assessing NAAQS compliance and population exposure and for performing model evaluations. The analysis is relatively straightforward, requiring only the site design values. The greater the design value, the higher the site rank. If more than one standard exists for a pollutant (e.g., annual and 24-hr average), monitors can be scored for each standard.

Example

This metric was one of five used in the 2000 National Analysis. The map shows the results for CO monitors. Sites in red record the highest CO concentrations and are the most valuable based on this metric. Sites in blue record the lowest values and are candidates for removal or repurposing.



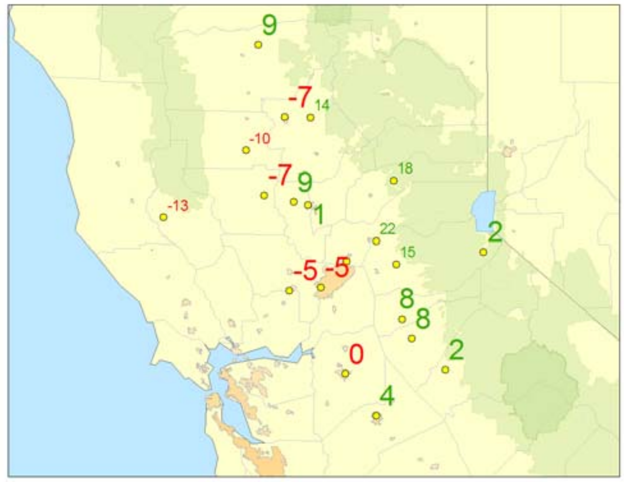
References

Schmidt M. (2001) Monitoring strategy: national analysis. Presented at the *Monitoring Strategy Workshop, Research Triangle Park, NC, October* by the U.S. Environmental Protection Agency. Available on the Internet at http://www.epa.gov/ttn/amtic/netamap.html.

Deviation from NAAQS

Overview

Sites that measure concentrations (design values) that are very close to the NAAQS exceedance threshold are ranked highest in this analysis. These sites may be considered more valuable for NAAQS compliance evaluation. Sites well above or below the threshold do not not provide as much information in terms of NAAQS compliance.



Ozone monitors in California and their deviation (ppb) from the maximum 8-hr NAAQS for a single day.

Type: Site-by-site analysis
Complexity: **
Size of network: any
Pollutants: Any with NAAQS or other standards

Objectives Assessed

- Regulatory compliance
- Forecasting assistance

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required			✓						
Helpful	✓			✓		✓			

Advantages

- Assesses monitor importance for determining NAAQS compliance

Disadvantages

- If design values vary from year to year, historical data should be included in the analysis
- Care is needed in interpreting absolute differences

Similar Analyses (Complexity)

Measured concentrations (**)
Removal bias (***)

Deviation from NAAQS

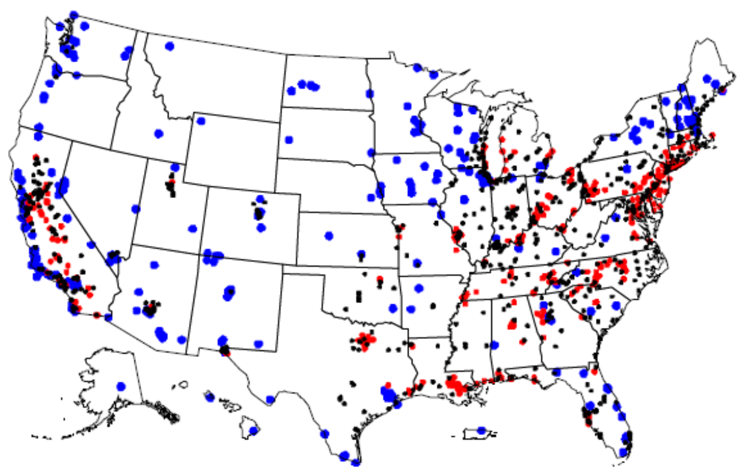
Analysis Goals

This technique contrasts the difference between the standard and actual measurements or design values. It is a simple way to assess a monitor's value for evaluating compliance. The design values for each pollutant should be calculated as they impact regulatory compliance. If a pollutant (e.g., annual and 24-hr average) has more than one standard, monitors can be scored for each standard. The absolute value of the difference between the measured design value and the standard can be used to score each monitor. Monitors with the smallest absolute difference will rank as most important. However, monitors that have higher design values than the standard (i.e., those in violation of the standard) may be considered more valuable from the standpoint of compliance and public health than those with design values lower than the standard, but with a similar absolute difference. Thus, absolute values of the difference can be ranked by peak concentration. It may be desirable to use more than one year of design values to look for consistency from year to year.

Example

Deviation from the NAAQS was one of five metrics used in the 2000 National Analysis. The analysis used one design value (1998–2000) and considered monitors above and below the standard equally. The map shows the results. Red circles denote sites that are nearest the standard, blue circles are those well above or below the standard, and black circles are in between.

Deviation from 1-hr O3 2nd Max NAAQS (98-00):



Interpretation

The red sites are ranked highest in this analysis. Depending on the network assessment objectives, the number of red-site monitors might be adjusted. Blue sites are candidates for removal or repurposing.

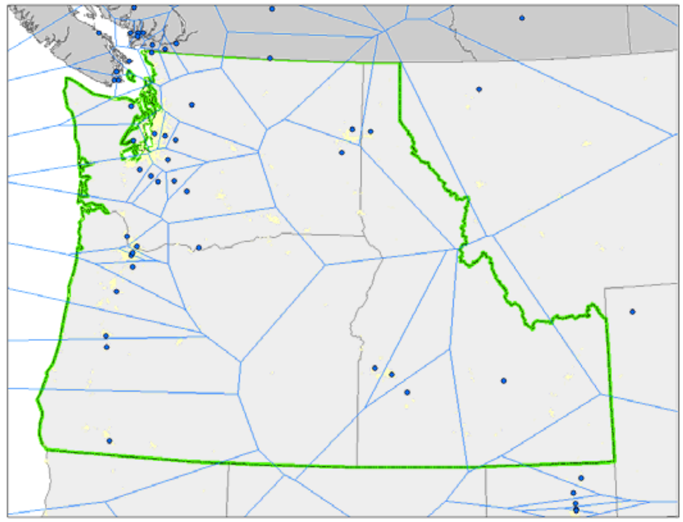
References

Schmidt M. (2001) Monitoring strategy: national analysis. Presented at the *Monitoring Strategy Workshop, Research Triangle Park, NC, October* by the U.S. Environmental Protection Agency. Available on the Internet at http://www.epa.gov/ttn/amtic/netamap.html.

Area Served

Overview

Sites are ranked based on their area of coverage. Sites that are used to represent a large area score high in this analysis. Area of coverage (area served) for a monitor can be determined using the Thiessen polygons technique. Each polygon consists of the points closer to one particular site than any other site. This technique gives the most weight to rural sites and those on the edges of urban areas or other monitor clusters.



Thiessan polygons showing the area served by ozone monitors (dots) in and around EPA Region 10.

Type: Site-by-site analysis
Complexity: **
Size of network: Moderate or larger
Pollutants: O_3, $PM_{2.5}$, SO_2, some toxics

Objectives Assessed

- Spatial coverage
- Interpolation
- Background concentration

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required	✓			✓					
Helpful									

Advantages

- Simple and quick to perform
- Gives weight to remote and urban boundary sites that are necessary for proper interpolation (e.g., for AIRNow maps)

Disadvantages

- Does not take into account topography or actual air basins
- Does not take into account population or emissions
- May artificially weight monitors at the edge of the analysis domain

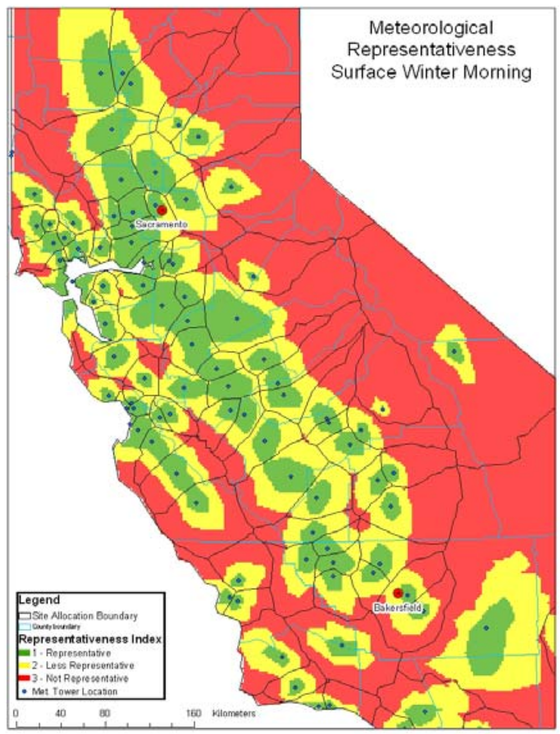
Similar Analyses (Complexity)

- Population served (***)
- Suitability modeling (****)

Area Served

Analysis Goals

Area served was one of five site-by-site criteria used in the national-scale network assessment. In the National Assessment, the "area served" metric was used as a proxy for the spatial coverage of each monitor. Theissen polygons (also called Voronoi diagrams) are applied as a standard technique in geography to assign a zone of influence or representativeness to the area around a given point. These polygons can be determined using a GIS package. Calculating Theissen polygons is one of the simplest quantitative methods for determining an area of representation around sites. However, it is not a true indication of which site is most representative in concentration to a given area. Meteorology (including pollutant transport), topography, and proximity to population or emission sources are not considered, so some areas assigned to a particular monitor may actually be better represented by a different monitor. More accurate determinations of representative monitors require a more sophisticated spatial analysis technique, such as suitability modeling, photochemical modeling, or parameter weighted distance.



Example

The map shows results of a study to determine zones of representativeness for meteorology towers in central California using a parameter-weighted distance technique. The method takes into account several factors to determine the "nearest" site: elevation, slope, time of day, season, height above ground, average wind speed, predominant wind direction, and geographic distance. The result is a zone of influence around each site that is more realistic than simple Theissan polygons, which only consider distance. In this map, the green areas are those that are best represented by the allocated tower for surface meteorological conditions during winter morning hours while red areas are not well represented by any of the existing measurements.

Interpretation

Regardless of the method for determining the boundaries of influence, the interpretation is the same. Sites with a greater area served are ranked higher than sites that only cover a small area. Sites that rank highly with this metric are valuable for interpolation, background concentration, and spatial coverage.

References

Knoderer C.A. and Raffuse S.M. (2004) CRPAQS surface and aloft meteorological representativeness (California Regional PM_{10}/$PM_{2.5}$ Air Quality Study Data Analysis Task 1.3). Web page prepared for the California Air Resources Board, Sacramento, CA, by Sonoma Technology, Inc., Petaluma, CA. Available on the Internet at http://www.sonomatechdata.com/crpaqsmetrep/ (STI-902324-2786).

O'Sullivan D. and Unwin D.J. (2003) *Geographic Information Analysis*, John Wiley & Sons, Inc., Hoboken, New Jersey.

U.S. Environmental Protection Agency (2001) National assessment of the existing criteria pollutant monitoring networks O_3, CO, NO_2, SO_2, Pb, PM_{10}, $PM_{2.5}$ - Part 1. Outputs from the National Network Assessment Introduction and Explanation, July 21. Available on the Internet at http://www.epa.gov/ttn/amtic/netamap.html.

Monitor-to-Monitor Correlation

Overview

Measured concentrations at one monitor are compared to concentrations at other monitors to determine if concentrations correlate temporally. Monitors with concentrations that correlate well (e.g., $r^2 > 0.75$) with concentrations at another monitor may be redundant. Conversely, a monitor with concentrations that do not correlate with other nearby monitored concentrations may be unique and have more value for spatial monitoring objectives. This analysis should be performed for each pollutant.

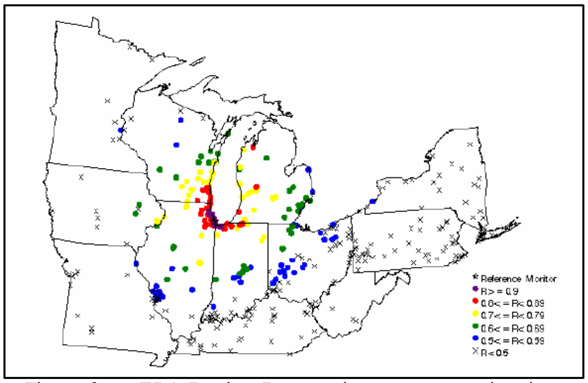


Figure from EPA Region 5 network assessment showing monitor-to-monitor correlation in and around the Chicago area.

Type: Site-by-site; Network optimization
Complexity: ** to ***
Size of network: large
Pollutants: O_3, $PM_{2.5}$, some toxics

Objectives Assessed

- Model evaluation
- Spatial coverage
- Interpolation

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required		✓	✓	✓					
Helpful	✓					✓	✓		

Advantages

- Gives measure of site's uniqueness and representativeness
- Useful for identifying redundant sites

Disadvantages

- Large data requirements
- Requires high data completeness
- Correlations are probably pollutant specific

Similar Analyses (Complexity)

- Measured concentrations (**)
- Principal Component Analysis (***)
- Removal Bias (***)

Monitor-to-Monitor Correlation

Analysis Goals

Determining the monitor-to-monitor correlation in a network requires at least two steps: (1) determining the temporal correlation between monitors through a regression analysis of concentrations; and (2) ranking the monitor's uniqueness. Step one can be accomplished most simply by calculating Pearson correlation coefficients (r^2) between each monitoring pair. Simple linear regressions can introduce error in the correlation coefficients, since they assume the ordinal axis has no error. Alternative methods include calculating Deming Regression or other types of correlation coefficients. In addition, choice of monitoring metrics may influence results (i.e., 1-hr peak ozone, every hour, 8-hr peak ozone, 24-hr average). Site pairs that have correlation coefficients with values near one are highly correlated and should be ranked lower than those with correlation coefficient values near zero. Sites that do not correlate well with other sites have unique temporal concentration variation relative to other sites and are likely to be important for assessing local emissions, transport, and spatial coverage. Conversely, those monitors that correlate with many other monitors may be redundant.

Example

This example shows a correlogram for ozone monitors located in the Chicago metropolitan area. Distance between monitors in kilometers is on the x-axis, and monitor-to-monitor correlation coefficients (r^2) are on the y-axis. The correlogram shows that ozone concentrations are highly correlated at most sites in Chicago with values above 0.8, and only decrease weakly as a function of distance.

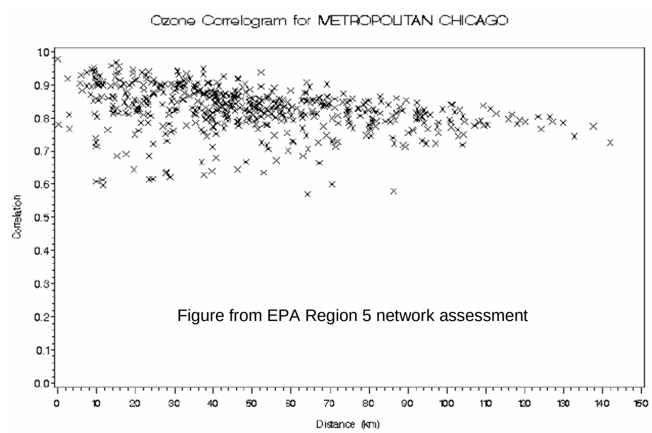


Figure from EPA Region 5 network assessment

This plot was created by calculating correlation coefficients and distance between sites.

Interpretation

This plot could be used to justify removing redundant sites, since concentrations correlate so well between most sites. Those monitor pairs with the lowest correlations (values around 0.6) would be rated as most important to retain.

References

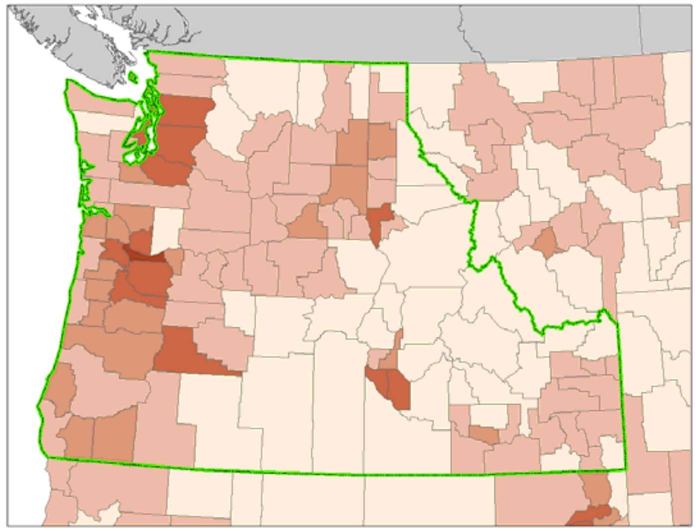
U.S. Environmental Protection Agency (2003) Region 5 network assessment. Presented at the *Air Monitoring & Quality Assurance Workshop, Atlanta, CA, September 9-11* by the U.S. Environmental Protection Agency, Region 5. Available on the Internet at <http://www.epa.gov/ttn/amtic/files/ambient/pm25/workshop/atlanta/r5netas.pdf> last accessed September 9, 2005.

Ito K., De Leon S., Thurston G.D., Nadas A., and Lippman M. (2005) Monitor-to-monitor temporal correlation of air pollution in the contiguous U.S. *J. Exposure Analy. Environ. Epidem.* **15**, 172-184.

Emission Inventory

Overview

Emission inventory data are used to find locations where emissions of pollutants of concern are concentrated. These locations can be compared to the current or proposed network. Does the network capture the areas of maximum emissions? This analysis can be scaled to various levels of complexity, depending on resources. The simplest version looks at county-level emissions of a single pollutant. More complex methods could use gridded emissions and/or species-weighted emissions, depending on their importance in producing the secondary pollutant(s) of cor



County level NO_x emission inventory for EPA Region 10. Darker shades represent greater emissions density.

Type: Bottom-up analysis
Complexity: ** to ****
Size of network: any
Pollutants: primary pollutants and secondary precursors

Objectives Assessed

- Emission reduction evaluation
- Maximum precursor location

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required								✓	
Helpful	✓			✓				Gridded and/or speciated	

Advantages

- Scaleable in complexity and spatial resolution
- Can find areas where primary pollutant concentrations will be high

Disadvantages

- Emission inventory data are not always current or may be incomplete or inaccurate
- Emission inventory quality varies by pollutant and source type
- More useful high resolution emission inventory data are not readily available and difficult to produce
- Does not consider transport

Similar Analyses (Complexity)

Site suitability modeling (****)

Emission Inventory

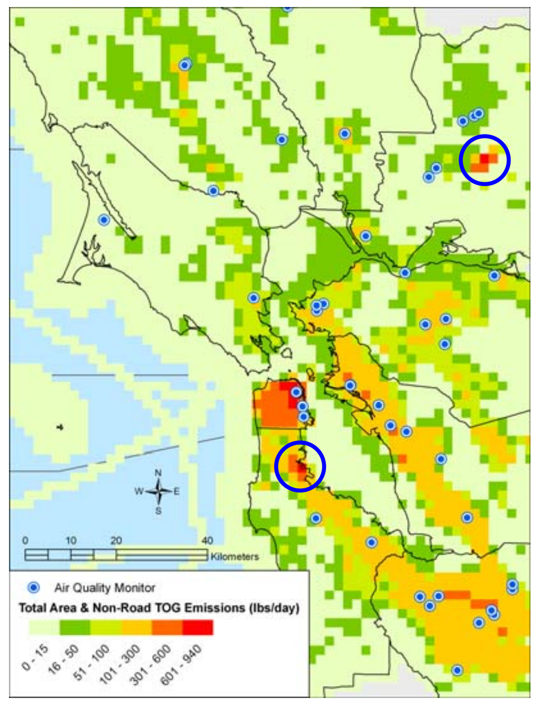
Analysis Goals

Emission inventory information is useful for determining locations of maximum emissions. At the simplest level, county-level emission data, such as the National Emission Inventory, can be compared with monitor locations. For measuring maximum precursor or primary emissions, monitors should be placed in those counties with maximum emission density (tons per year per square mile). More refined site placement decisions can be considered with more refined emission inventory data or wind data to indicate the up- and downwind directions. State and local air quality agencies can supply gridded emission inventories, which will depict more focused areas for measuring maximum precursor or primary emissions. Speciated emissions inventory data can also be used. The process of disaggregating inventory pollutants into individual chemical species components or groups of species will help determine placement of monitors that have pollutant-specific monitoring objectives.

Example

This example in and around the San Francisco Bay Area was created in ESRI ArcGIS 9.1 using the following steps:

1. Acquire a gridded emission inventory for the greater San Francisco Bay Area
2. Overlay an existing monitor network on the gridded inventory
3. Determine the estimated emissions amount at each monitor location based on the grid cell it falls within
4. For areas with high estimated emissions values, calculate the distance to the closest monitor location



Interpretation

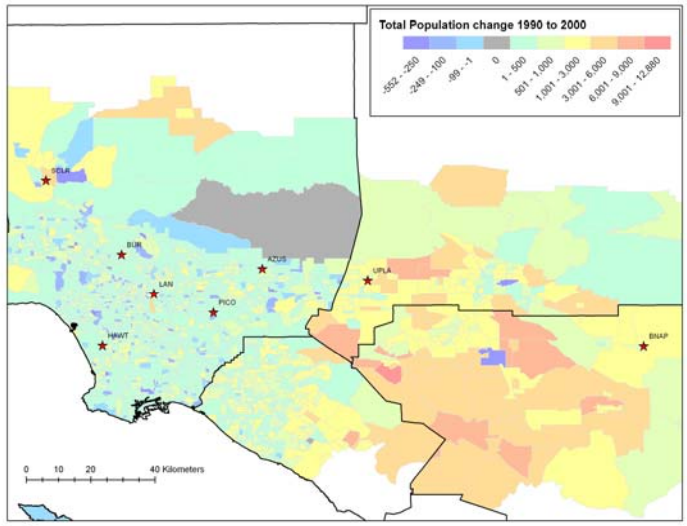
The table at right is an extract of the analysis example for the San Francisco Bay Area. The high emission estimates listed by grid cell coincide with a distance value to the closest monitor location. A zero distance means a monitor is located within that grid cell. The table shows that the grid cell containing the largest amount of estimated emissions has a monitor over 10 kilometers away. The two blue circles on the map show areas of high emission density with no current monitors. These areas may be good candidates for future monitoring sites.

Cell ID ; emissions (lbs/day)	Distance to closest monitor (kilometers)
4850 ; 936	10.7
1099 ; 788	0
1323 ; 777	1.4
3395 ; 664	2.1
745 ; 655	11.5
4021 ; 627	3.1
5223 ; 585	2.7
788 ; 565	8.5

Population Change

Overview

High rates of population increase are associated with increased potential emissions activity. Sites are ranked based on population increase in the area of representation. Area of representation can be determined using the Thiessen polygons technique or a more sophisticated method (see Area Served). The total population change at the census-tract or block-group level that falls within the area of coverage of a monitor is assigned to that monitor. This technique gives most weight to sites in areas with high rates of population growth and large areas of representation.



1990 to 2000 population change in and around Los Angeles.

Type: Site-by-site analysis; bottom up
Complexity: ***
Size of network: any
Pollutants: O_3, $PM_{2.5}$, SO_2, some toxics

Objectives Assessed

- Population exposure
- Environmental justice
- Maximum precursor location

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required	✓			✓	✓	✓			
Helpful									Demographics

Advantages

- Assesses site importance for population exposure, an important regulatory goal
- Flexible (a few possible methods)
- Helpful for determining where monitoring may be required in the future
- Helps identify monitors near which emissions may have substantially changed

Disadvantages

- Does not take into account topography or actual air basins (using basic method)
- Highly resolved population data may be difficult to work with
- Changing census boundaries make it difficult to compare populated areas over time

Similar Analyses (Complexity)

- Area served (**)
- Population served (***)
- Suitability modeling (****)

Population Change

Analysis Goals

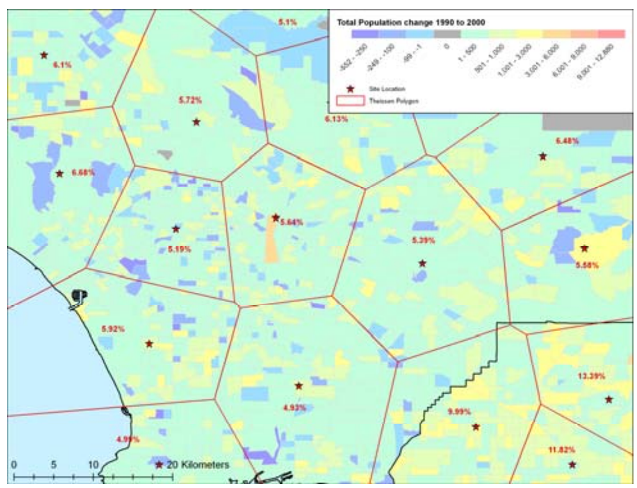
Determining the population change near a particular monitor requires two steps: (1) identify the area of responsibility for each monitor; and (2) determine the percent change in population within each area of responsibility. Step 1 can be done most simply using the Theissen polygons technique; however, a more sophisticated method that takes into account distance, meteorology, topography, etc. can also be applied (see Area Served). Step 2 can be performed using U.S. Census population data at a variety of geographic levels (i.e., census block group, census tract). However, because census boundaries change over time, it is difficult and time-intensive to link localized census boundary data. The link between census boundary files is necessary to join the comparison population values and find an accurate percent change in population. One way to accomplish this is by gridding both data sets to a common grid scale.

Sites that score high with this metric are important for assessing population exposure and tracking future emissions growth. The population change method can also be applied to assess the importance of monitors from an environmental justice perspective. The technique is the same, except population changes of specific groups (e.g., low income or minority) are calculated instead of total population. Population change can also be applied as a bottom-up technique. Using the census data, areas of rapid growth can be located and considered as potential locations for new monitors.

Example

This example in and around the Los Angeles, California, area was created in ESRI ArcGIS 9.1 using the following steps:

1. Create Theissen polygon coverage of monitoring sites
2. Link the 1990 and 2000 census tract polygons by tract ID in order to get total change in population by census tract
3. Convert census tract polygons to centroid points
4. Calculate the percent change in population for each monitoring area by spatially joining Theissen polygons to census tract centroids



Interpretation

The table at right is an extract of the analysis example for Los Angeles. The area around site location 4 has seen a 13% increase in population and has, therefore, grown in importance for monitoring population exposure.

Site Location	% Population Change 1990 to 2000
1	5%
2	12%
3	10%
4	13%
5	5%
6	6%
7	5%
8	5%

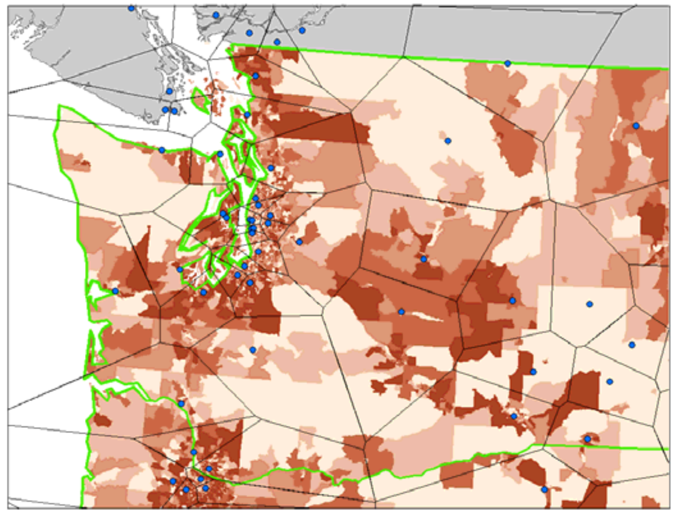
References

U.S. Census Bureau, Census 2000; 1990 Census, Population and Housing Unit Counts, United States. Available at <http://www.census.gov/>)

Population Served

Overview

Large populations are associated with high emissions. Sites are ranked based on the number of people they represent. Area of representation can be determined using the Thiessen polygons technique or a more sophisticated method (see Area Served). Populations at the census-tract or block-group level that fall within the area of representation of a monitor are assigned to that monitor. This technique gives the most weight to sites that are in areas of high population and have large areas of representation.



Population density and ozone monitor areas of representation in western Washington. Darker colors represent greater population.

Type: Site-by-site analysis
Complexity: ***
Size of network: Moderate or larger
Pollutants: O_3, $PM_{2.5}$, SO_2, some toxics

Objectives Assessed

- Population exposure
- Environmental justice

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required	✓			✓	✓				
Helpful									Demographics

Advantages

- Assesses site importance for population exposure, an important regulatory goal
- Flexible (a few possible methods)

Disadvantages

- Does not take into account topography or actual air basins (using basic method)
- Highly resolved population data may be difficult to work with

Similar Analyses (Complexity)

- Area served (**)
- Counties served (**)
- Population change (***)
- Suitability modeling (****)

Population Served

Analysis Goals

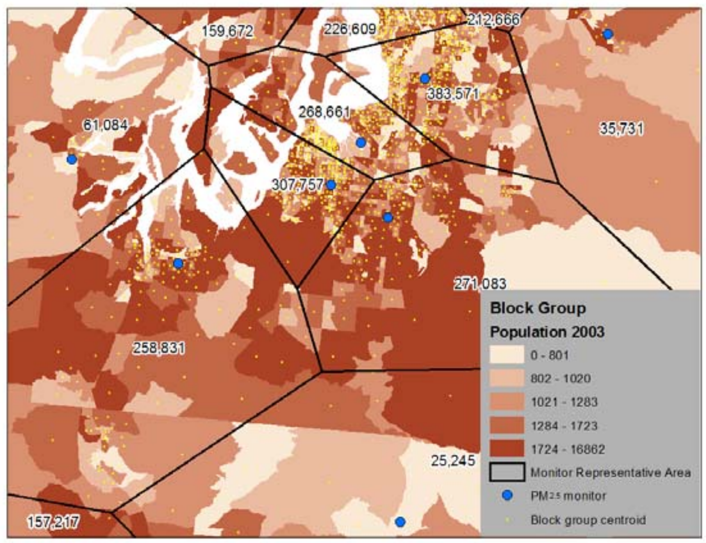
Calculating the population served by a particular monitor requires two steps: (1) determine the area of representation for each monitor; and (2) determine the population within each area of representation. Step 1 can be performed most simply using the Theissen polygons technique; however, a more sophisticated method that takes into account distance, meteorology, topography, etc. could also be applied (see Area Served). Sites that score high with this metric are important for assessing population exposure. This technique was one of five site-by-site criteria used in the national-scale network assessment. Theissen polygons (also called Voronoi diagrams) are applied as a standard technique in geography to assign a zone of influence or representativeness to the area around a given point.

The "population served" method can also be applied to assess the importance of monitors from an environmental justice perspective. The technique is the same, except populations of specific groups (e.g., low income or disadvantaged) are used instead of total population.

Example

This example in and around the Seattle, Washington, area was created in ESRI ArcGIS 9.1 using the following steps:

1. Create Theissen polygon coverage of $PM_{2.5}$ monitoring sites
2. Convert census block group polygons (available on ESRI data CDs) to centroid points
3. Sum population in each monitoring area by spatially joining Theissen polygons to block group centroids



Interpretation

The table at right is an extract of the analysis example for Washington State. Note that the population served varies by two orders of magnitude. The actual population values could be used to weight the sites, or they could simply be ranked. If the population values are used, the highly populated monitor sites will be given much greater weight than the sparsely populated monitor sites. This method could also be used within a network optimization assessment. For each network scenario, an average population served can be calculated. Scenarios with a lower average population served cover fewer persons per monitor, which may be less desirable.

AIRS Code	Population Served
530630016	423,089
530332004	383,571
530110013	379,893
530610005	349,160
530750003	32,633
530210002	28,538
530330037	25,245
530750006	12,363
530130001	9,092
530010003	8,961
530750005	2,392

References

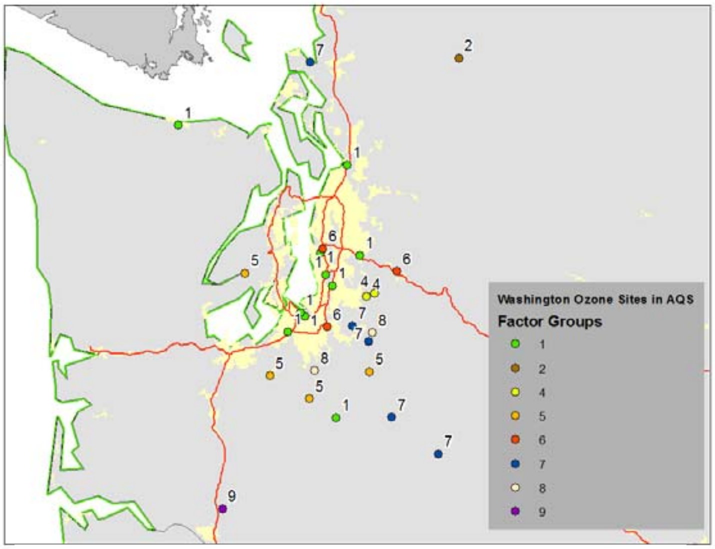
U.S. Environmental Protection Agency (2001) National assessment of the existing criteria pollutant monitoring networks O_3, CO, NO_2, SO_2, Pb, PM_{10}, $PM_{2.5}$ - Part 1. Outputs from the National Network Assessment Introduction and Explanation, July 21. Available on the Internet at http://www.epa.gov/ttn/amtic/netamap.html.

O'Sullivan D. and Unwin D.J. (2003) *Geographic Information Analysis*, John Wiley & Sons, Inc., Hoboken, New Jersey.

Principal Component Analysis

Overview

Principal Component Analysis (PCA) can be applied to find monitoring sites that have a pattern of variability similar to other monitoring sites. PCA assigns each monitor to a group of monitors at which pollutant concentrations behave similarly to each other. This analysis can be useful for finding redundancy in the network. It is also useful in selecting sites for other analyses (e.g., source apportionment).



Example of resulting factor groups for ozone monitoring sites in the Seattle area.

Type: Network Optimization
Complexity: ***
Size of network: large
Pollutants: O_3, $PM_{2.5}$, SO_2, toxics

Objectives Assessed

- Background concentration
- Forecasting assistance

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required		✓	✓	✓					
Helpful	✓					✓	✓		

Advantages

- Can identify potentially redundant monitors
- Highlights spatial trends in data that help identify hot spots and large sources
- Useful for site selection for other investigatory analyses
- Identifies areas in similar air basins

Disadvantages

- Requires analyst skill to avoid over-interpretation
- Groups monitors by variability, not by concentration
- Some monitors may appear in multiple groups
- Requires high data completeness and lots of data

Similar Analyses (Complexity)

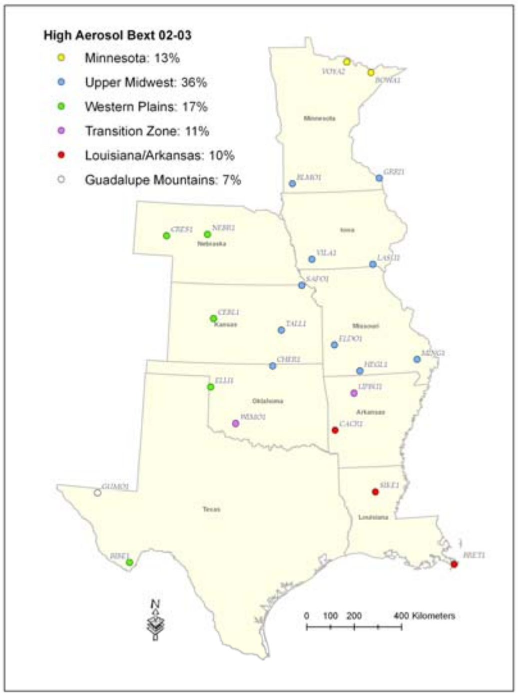
Monitor-to-monitor correlation (** to ***)
Removal bias (***)

Principal Component Analysis

Analysis Goals

PCA is a useful tool for examining possible monitor redundancies. PCA identifies recurring and independent signals within large and noisy data sets such as ambient data (Eder et al., 1993). The results can be used to identify groups of sites with similar variance in measured concentrations. PCA is commonly available in statistical software packages. Hourly or daily samples (the more the better) with high data completeness at each site are required to perform the analysis.

Example



The example comes from an analysis of visibility measurements at Class I areas in the Central Regional Air Planning Association (CENRAP). Each color represents an identified cluster of sites that have similar variance patterns in visibility. Similar techniques have been applied to ozone in rural sites in the eastern United States (Eder et al., 1993; Lehman et al., 2004).

For its 2002 network assessment, EPA Region 5 performed positive matrix factorization (PMF) on ozone monitors. PMF is a more complex analysis that achieves similar goals.

Interpretation

The direct outputs from PCA or other factor analysis tools are not site groupings. Rather, they are principal components that describe a percentage of the concentration variance at a particular site. Sometimes, a given site may be in multiple principal components (factors), which can indicate a site that is in a transition zone between factors. Therefore, the results require interpretation to assign a specific monitor to a particular group or to understand the "transition zones" in the network. The groupings are useful to select sites for additional analyses, assess zones of influence for a given pollutant, and identify possible redundant sites.

References

Eder B.K., Davis J.M., and Bloomfield P. (1993) A characterization of the spatiotemporal variability of non-urban ozone concentrations over the eastern United States. *Atmos. Environ.* **27A**, 2645-2668.

Lehman J., Swinton K., Bortnick S., Hamilton C., Baldridge E., Eder B., and Cox B. (2004) Spatio-temporal characterization of tropospheric ozone across the eastern United States. *Atmos. Environ.* **38**, 4357-4369.

Sullivan D.C., Hafner H.R., Brown S.G., MacDonald C.P., Raffuse S.M., Penfold B.M., and Roberts P.T. (2005) Analyses of the causes of haze for the Central States (phase II) summary of findings. Executive summary prepared for the Central States Regional Air Planning Association by Sonoma Technology, Inc., Petaluma, CA, STI-904780.08-2754-ES, August.

U.S. Environmental Protection Agency (2003) Region 5 network assessment. Presented at the *Air Monitoring & Quality Assurance Workshop, Atlanta, CA, September 9-11* by the U.S. Environmental Protection Agency, Region 5. Available on the Internet at http://www.epa.gov/ttn/amtic/files/ambient/pm25/workshop/atlanta/r5netas.pdf last accessed September 9, 2005.

Removal Bias

Overview

Measured values are interpolated across the domain using the entire network. Sites are then systematically removed and the interpolation is repeated. The absolute difference between the concentration measured at a site and the concentration predicted by interpolation with the site removed is the site's removal bias. The greater the bias, the more important the site is for interpolation. This analysis can also be performed on groupings of sites to test various site removal scenarios.

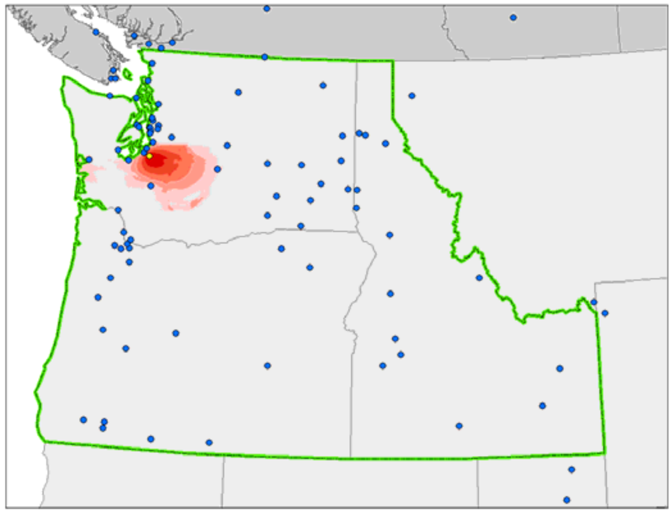


Figure showing the removal bias of a single monitoring site in EPA Region 10. The darker shade represents greater bias.

Type: Site-by-site; Network optimization
Complexity: ***
Size of network: large
Pollutants: O_3, $PM_{2.5}$, SO_2, some toxics

Objectives Assessed

- Interpolation
- Spatial coverage
- NAAQS compliance
- Background concentration
- Model evaluation

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required	✓		✓	✓					
Helpful		✓							

Advantages

- Gives measure of site's importance for several objectives
- Useful for site-by-site ranking and network optimization

Disadvantages

- Requires geostatistical tools
- Does not account for geographic features
- Most useful for pollutants with large networks

Similar Analyses (Complexity)

- Monitor-to-monitor correlation (**)
- Principal Component Analysis (***)

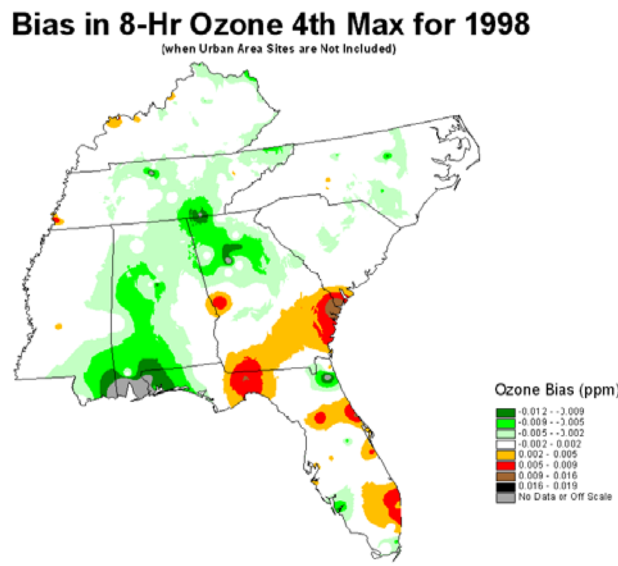
Removal Bias

Analysis Goals

Removal bias is a sensitivity analysis to determine how important a particular monitor (or set of monitors) is for interpolating concentrations across the domain. Variations of this method were performed in the National Analysis, as well as the draft assessments for EPA Regions 3 and 4. The basic method is to compare interpolations with and without data from specific monitors to determine either the bias or uncertainty that results from the removal of those monitors. Greater bias or uncertainty indicates a more important site for developing interpolations to represent concentrations across the domain. Those sites with low bias may be providing information that is redundant. With a base concentration field across the entire domain (developed through photochemical modeling), hypothetical monitors can also be tested.

Example

For the National Analysis, a site-by-site approach was used. That is, each site was removed individually and the resulting uncertainty at the site was calculated. Region 4 applied a network optimization technique, removing certain classes of sites (e.g., rural, urban core) and calculating interpolation bias. The image at right is from the Region 4 assessment. It shows the bias in 8-hr ozone when all urban sites are removed: positive bias is shown in red and negative bias in green.



Interpretation

It is perhaps counterintuitive that removing all urban sites would produce a positive bias in concentrations of ozone. This is likely because 8-hr ozone concentrations are often at maximum downwind from the areas of maximum precursor emissions (urban areas).

When looking at individual contributions to bias or uncertainty, as in the National Analysis, it is important to avoid over-interpretation. For example, clustered sites may all have low individual biases because of their redundancy and may all be candidates for removal. However, removing all of those sites would potentially create a large bias in the area. If a suite of monitors are targeted for removal, it would be useful to perform a bias analysis on the resulting network to ensure that the combined effects of removal are acceptable.

References

Cimorelli A.J., Chow A.H., Stahl C.H., Lohman D., Ammentorp E., Knapp R., and Erdman T. (2003) Region III ozone network reassessment. Presented at the *Air Monitoring & Quality Assurance Workshop, Atlanta, GA, September 9-11* by the U.S. Environmental Protection Agency, Region 3, Philadelphia, PA. Available on the Internet at http://www.epa.gov/ttn/amtic/files/ambient/pm25/workshop/atlanta/r3netas.pdf last accessed September 9, 2005.

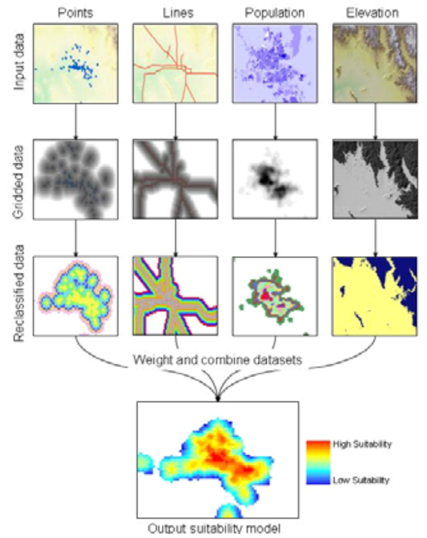
Schmidt M. (2001) Monitoring strategy: national analysis. Presented at the *Monitoring Strategy Workshop, Research Triangle Park, NC, October* by the U.S. Environmental Protection Agency. Available on the Internet at http://www.epa.gov/ttn/amtic/netamap.html.

U.S. Environmental Protection Agency (2002) Assessment of the ambient air monitoring networks. Draft report prepared for the U.S. Environmental Protection Agency, Research Triangle Park, NC, by the U.S. Environmental Protection Agency, Region 4, October. Available on the Internet at http://www.epa.gov/ttn/amtic/files/ambient/pm25/workshop/atlanta/r4netas.pdf last accessed September 9, 2005.

Suitability Modeling

Overview

Suitability modeling is a method for identifying suitable monitoring locations based on specific criteria. Geographic map layers representing important criteria, such as emissions source influence, proximity to populated places, urban or rural land use, and site accessibility can be compiled and merged to develop a composite map representing the combination of important criteria for a defined area. Furthermore, each map layer input can be assigned a weighting factor based on the relative importance of each layer in the overall suitability model. The results provide the best locations to site monitors based on the input criteria.



Suitability model conceptual diagram. Input feature data are converted to gridded surfaces, classified to a common scale, weighted, and combined to form the output model.

Type: Bottom-up analysis
Complexity: ****
Size of network: any
Pollutants: any

Objectives Assessed

- Population exposure
- Environmental justice
- Source-oriented monitoring
- Model evaluation
- Maximum concentration location
- Background concentration
- Transport/border characterization

Resources

	Tools		Data						
	GIS	Statistical Software	Concentrations	Site Locations	Population	Historical Data	Site Information	Emission Inventory	Other
Required	✓			✓	✓		✓	✓	Demographics
Helpful			✓						Meteorology

Advantages

- Assesses site importance for population exposure—an important regulatory goal
- Flexible (able to run several model scenarios)
- Does not require ambient data
- Graphic results are useful to a broad audience

Disadvantages

- Time-intensive
- Weighting scheme is subjective; analysis is iterative
- Requires skilled GIS analyst
- GIS data layers can be difficult and costly to acquire

Similar Analyses (Complexity)

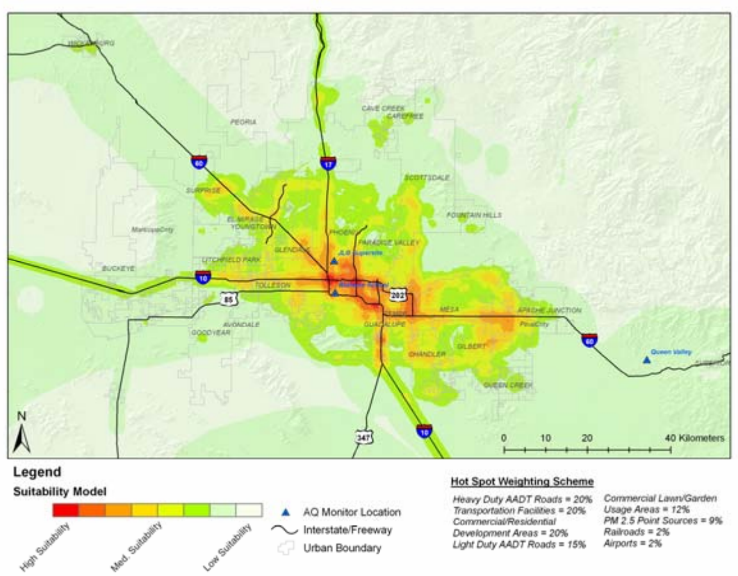
Area served (**)
County served (**)
Population served (***)
Population change (***)

Suitability Modeling

Analysis Goal

The first step of a suitability analysis involves selecting criteria that can address monitoring objectives. The second step of the analysis is to acquire and process the spatial data for the suitability model within a GIS. The third and last step is to develop and run a suitability model for different model scenarios (see analysis approach figure below).

Example



In this example, suitability modeling was used to determine candidate sites for monitoring diesel particulate matter (DPM) in and around Phoenix, Arizona. Because we are interested in identifying locations where emissions for a particular pollutant are likely to be high, we must be able to spatially characterize the distribution of emissions for each major pollutant source category.

The example was created using ESRI ArcGIS 9.1 (Spatial Analysis extension) using the following steps:

1. Assess an emission inventory to determine the predominant sources of DPM in the region and determine the best available data to represent the spatial pattern of the identified emissions sources in the Phoenix region.
2. Acquire and process the spatial data (map layers) required for the analysis. For example, a map of roadways and associated traffic volumes for heavy- and light-duty vehicles were used to characterize the spatial distribution of emissions from on-road mobile sources.
3. Develop and run the suitability model for different model scenarios. Three model scenarios were defined to examine the spatial distribution of DPM emissions: (1) development of a composite map to represent the spatial distribution and density of DPM emissions based on the locations of DPM sources (hot spots), (2) proximity of total population to DPM sources, and (3) proximity of sensitive population groups to DPM sources (see the figure above).

Interpretation

Existing monitor locations, not originally located to investigate DPM, were suitable. Other locations in this fast-growing area were identified that would be suitable for assessment of DPM impacts on the population.

References

Hafner H.R., Penfold B.M., and Brown S.G. (2005) Using spatial analysis techniques to select monitoring locations. Presentation at the *U.S. Environmental Protection Agency's 2005 National Air Quality Conference: Quality of Air Means Quality of Life, San Francisco, CA, February 12-13* (STI-2645).

4. RECOMMENDATIONS

Network assessment facilitates developing an optimal balance between scientific quality, protection of public and environmental health and welfare, and available resources. It is a tool for identifying opportunities

- to redistribute resources to valued programs from low-priority or low-benefit ones;
- to create "found money" for programs previously thought to be unaffordable;
- to extract more value from existing networks; and
- to fully leverage the value EPA's or other agencies' existing networks.

Before beginning a network assessment, the purposes of the monitoring network—i.e., the network's mission (e.g., establish regulatory compliance, further scientific understanding)—should be established or carefully revisited and prioritized. With the network's purposes and priorities in mind, users of this guidance document may perform the analyses described singly or in combination to design a technical network assessment suitable for their circumstances. Site-by-site comparisons help identify monitoring sites within an existing network that are most or least valuable relative to the purposes of the network. Bottom-up analyses yield appraisals of existing monitoring sites' value relative to their optimal placement. Network optimization analyses are particularly useful when considering alternative scenarios for network design.

This guidance document addresses specific technical elements of the overall framework for network assessments that is discussed in the draft NAAMS (U.S. Environmental Protection Agency, 2005)—specifically, statistical and objective situation analyses. However, the NAAMS recognizes other key elements of network assessment, such as *subjective* situational analyses, preparation of regional descriptions or network histories, and solicitation of input from state and local agencies or stakeholders. Further, the NAAMS acknowledges the importance of considering non-technical factors, such as political or justice-related issues. Therefore, this guidance document represents a starting point for the development of further guidance for network assessments. It is designed to be flexible and expandable with additional types of technical analyses and examples as techniques are improved, enhanced, and more broadly applied. In addition, the development of additional guidance covering other key elements of network assessments and non-technical considerations are areas for further work.

5. REFERENCES

Cimorelli A.J., Chow A.H., Stahl C.H., Lohman D., Ammentorp E., Knapp R., and Erdman T. (2003) Region III ozone network reassessment. Presented at the *Air Monitoring & Quality Assurance Workshop, Atlanta, GA, September 9-11* by the U.S. Environmental Protection Agency, Region 3, Philadelphia, PA. Available on the Internet at <http://www.epa.gov/ttn/amtic/files/ambient/pm25/workshop/atlanta/r3netas.pdf> last accessed September 9, 2005.

Clean Air Scientific Advisory Committee and National Ambient Air Monitoring Strategy Subcommittee (2003) Summary minutes of the Clean Air Scientific Advisory Committee (CASAC) National Ambient Air Monitoring Strategy (NAAMS) Subcommittee pre-meeting public teleconference. Summary minutes prepared for the Clean Air Scientific Advisory Committee and National Ambient Air Monitoring Strategy Subcommittee Teleconference, Research Triangle Park, NC, July.

Eder B.K., Davis J.M., and Bloomfield P. (1993) A characterization of the spatiotemporal variability of non-urban ozone concentrations over the eastern United States. *Atmos. Environ.* **27A**, 2645-2668.

Hafner H.R., Penfold B.M., and Brown S.G. (2005) Using spatial analysis techniques to select monitoring locations. Presentation at the *U.S. Environmental Protection Agency's 2005 National Air Quality Conference: Quality of Air Means Quality of Life, San Francisco, CA, February 12-13* (STI-2645).

Henderson R. (2005) Clean Air Scientific Advisory Committee advisory on implementation aspects of the U.S. Environmental Protection Agency's final draft National Air Monitoring Strategy. Letter written to Stephen L. Johnson, Acting Administrator for the U.S. Environmental Protection Agency, Washington, DC, by the Clean Air Scientific Advisory Committee, April.

Hopke P.K. (2003) Clean Air Scientific Advisory Committee review of the U.S. Environmental Protection Agency's National Ambient Air Monitory Strategy. Letter written to Michael O. Leavitt, Administrator for the U.S. Environmental Protection Agency, Washington, DC, by the Clean Air Scientific Advisory Committee, November.

Ito K., De Leon S., Thurston G.D., Nadas A., and Lippman M. (2005) Monitor-to-monitor temporal correlation of air pollution in the contiguous U.S. *J. Exposure Analy. Environ. Epidem.* **15**, 172-184.

Knoderer C.A. and Raffuse S.M. (2004) CRPAQS surface and aloft meteorological representativeness (California Regional $PM_{10}/PM_{2.5}$ Air Quality Study Data Analysis Task 1.3). Web page prepared for the California Air Resources Board, Sacramento, CA, by Sonoma Technology, Inc., Petaluma, CA. Available on the Internet at <http://www.sonomatechdata.com/crpaqsmetrep/> (STI-902324-2786).

Lehman J., Swinton K., Bortnick S., Hamilton C., Baldridge E., Eder B., and Cox B. (2004) Spatio-temporal characterization of tropospheric ozone across the eastern United States. *Atmos. Environ.* **38**, 4357-4369.

Main H.H. and Roberts P.T. (2001) Recommendations for the PAMS network in the Northeast and Mid-Atlantic States. Report prepared for NESCAUM, Boston, MA, by Sonoma Technology, Inc., Petaluma, CA, STI-900860-2067-FR, June.

O'Sullivan D. and Unwin D.J. (2003) *Geographic Information Analysis*, John Wiley & Sons, Inc., Hoboken, New Jersey.

Park S.S., Harrison D., Pancras J.P., and Ondov J.M. (2005) Highly time-resolved organic and elemental carbon measurements at the Baltimore Supersite in 2002. *J. Geophys. Res.* **110** (D07S06) (doi:10.1029/2004JD004610).

Scheffe R. (2003) National Ambient Air Monitoring Strategy: a comprehensive re-examination and re-configuration of air monitoring networks. Presented at the *Clean Air Scientific Advisory Committee (CASAC) National Ambient Air Monitoring Strategy (NAAMS) Subcommittee Teleconference, Research Triangle Park, NC, July*.

Schmidt M. (2001) Monitoring strategy: national analysis. Presented at the *Monitoring Strategy Workshop, Research Triangle Park, NC, October* by the U.S. Environmental Protection Agency. Available on the Internet at http://www.epa.gov/ttn/amtic/netamap.html.

Stahl C.H., Cimorelli A.J., and Chow A.H. (2002) A new approach to environmental decision analysis: Multi-Criteria Integrated Resource Assessment (MIRA). *Bulletin of Science, Technology, & Society* **22** (6), 443-459.

Sullivan D.C., Hafner H.R., Brown S.G., MacDonald C.P., Raffuse S.M., Penfold B.M., and Roberts P.T. (2005) Analyses of the causes of haze for the Central States (phase II) summary of findings. Executive summary prepared for the Central States Regional Air Planning Association by Sonoma Technology, Inc., Petaluma, CA, STI-904780.08-2754-ES, August.

U.S. Environmental Protection Agency (1998) SLAMS/NAMS/PAMS network review guidance. Revised final report prepared for the U.S. Environmental Protection Agency, Emissions, Monitoring, and Analysis Division, Research Triangle Park, NC, March.

U.S. Environmental Protection Agency (2001) National assessment of the existing criteria pollutant monitoring networks O_3, CO, NO_2, SO_2, Pb, PM_{10}, $PM_{2.5}$ - Part 1. Outputs from the National Network Assessment Introduction and Explanation, July 21. Available on the Internet at http://www.epa.gov/ttn/amtic/netamap.html.

U.S. Environmental Protection Agency (2002) Assessment of the ambient air monitoring networks. Draft report prepared for the U.S. Environmental Protection Agency, Research Triangle Park, NC, by the U.S. Environmental Protection Agency, Region 4, October. Available on the Internet at <http://www.epa.gov/ttn/amtic/files/ambient/pm25/workshop/atlanta/r4netas.pdf> last accessed September 9, 2005.

U.S. Environmental Protection Agency (2003a) AirNow web site. Available on the Internet at <http://www.airnow.gov> last accessed January 17, 2005.

U.S. Environmental Protection Agency (2003b) Air Monitoring & Quality Assurance Workshop - September 2003. *Air Monitoring & Quality Assurance Workshop, Atlanta, GA, September 9-11, 2003*. Available on the Internet at <http://www.epa.gov/ttn/amtic/pmatla.html> last accessed September 9, 2005.

U.S. Environmental Protection Agency (2003c) Region 5 network assessment. Presented at the *Air Monitoring & Quality Assurance Workshop, Atlanta, CA, September 9-11* by the U.S. Environmental Protection Agency, Region 5. Available on the Internet at <http://www.epa.gov/ttn/amtic/files/ambient/pm25/workshop/atlanta/r5netas.pdf> last accessed September 9, 2005.

U.S. Environmental Protection Agency (2005) Draft National Ambient Air Monitoring Strategy. Draft report by the U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, December.

APPENDIX A

EXAMPLE APPLICATIONS

Several network assessments and other projects related to network assessment have already been executed for air quality monitoring networks. Here we present a few examples, covering a range of scales, objectives, and available resources.

A.1 NATIONAL ASSESSMENT

A national assessment was performed in 2000 on the criteria pollutant monitoring network. Its goal was to provide broad directional recommendations on a national level and act as a guide for more focused regional (and local) assessments. The National Assessment (U.S. Environmental Protection Agency, 2001) utilized five site-by-site analysis metrics and several weighting schemes to rank individual monitors. Each pollutant was considered separately. The five metrics used were (in the terms of this document) area served, population served, measured concentrations, deviation from NAAQS, and uncertainty on removal. Divestment opportunities were highlighted using the ranked monitors. The National Assessment is available at http://www.epa.gov/ttn/amtic/netamap.html.

A.2 REGIONAL ASSESSMENTS

In fiscal year 2003, each of the ten EPA regions began their own network assessments, building from the 2000 National Assessment. These assessments varied greatly in their methods and depth, partly motivating this document. The approaches taken by three of the regions are highlighted below.

A.2.1 Region 3 Ozone Network Assessment

EPA Region 3 employed a network optimization technique for its network assessment. The technique was based on an iterative analytical decision making processed called Multi-Criteria Integrated Resource Assessment (MIRA) (Cimorelli et al., 2003; Stahl et al., 2002). In brief, the technique started with several possible network configurations and ranked them using 40 decision criteria organized hierarchically; the four primary level criteria are trends impact, costs, air quality, and personnel impact. The "air quality" criterion was the ability of the network to capture and properly interpolate concentrations from a base-case scenario developed with photochemical modeling. MIRA incorporates stakeholder feedback and participation throughout the process, and network configurations, design criteria, and weighting schemes were modified as learning proceeded.

A.2.2 Region 4 Network Assessment

EPA Region 4 utilized EPA monitoring re-engineering guidance (U.S. Environmental Protection Agency, 1998) for all criteria monitors except ozone and $PM_{2.5}$. The re-engineering guidance suggests monitors that do not exceed 60% of the NAAQS are candidates for termination. From this baseline, Region 4 worked with state and local agencies to determine which candidate monitors were of low value or redundant and which monitors provided useful research information or satisfied regulatory requirements.

Because none of the ozone monitors and only one of the $PM_{2.5}$ monitors in Region 4 were below the 60% threshold, additional geospatial analyses were performed using the National Assessment (U.S. Environmental Protection Agency, 2001) as a guide. Region 4 used the removal bias technique to determine the effects of removing certain classes of monitors (e.g., urban core, downwind, upwind).

A.2.3 Region 5 $PM_{2.5}$ Network Assessment

The EPA Region 5 $PM_{2.5}$ network assessment process was organized by the Lake Michigan Air Directors Consortium (LADCO) (U.S. Environmental Protection Agency, 2003c). It is a site-by-site analysis, similar to the National Assessment, and considered four metrics: measured concentrations, monitor-to-monitor correlation, population change, and monitor density. Rather than weighting each metric and developing a combined score, the metrics were considered in a stepwise fashion. Sites were first ranked only on the most important metric (monitor density), the highest scoring monitors (i.e., those farthest from other monitors) were then eliminated from consideration for removal. The remaining monitors were ranked based on the next most important metric and so on. The monitors in the final list were then considered individually for possible elimination.

A.3 PHOTOCHEMICAL ASSESSMENT MONITORING STATIONS (PAMS) NETWORK ASSESSMENT

In 2001, the PAMS network was assessed with the goal of balancing and redirecting resources to meet evolving program objectives (Main and Roberts, 2001). Starting with the existing network, the analysis identified the minimum type and number of observations required to satisfy PAMS goals, determined what monitors met the those goals, developed recommendations for eliminating monitors that were not required, and identified ways to further enhance the PAMS program in the long-term with the resources saved. To determine monitors that could be eliminated, the study looked at site pairs in close proximity and performed several statistical data analysis techniques to determine similarity, including medians, interquartile ranges, confidence intervals, and p-values. They also found that some sites were designated as types (upwind, maximum ozone, etc.) that did not match the data.

A.4 PHOENIX DIESEL PARTICULATE MATTER SITE SUITABILITY MODELING

Diesel particulate matter (DPM) is an issue of increasing concern for protecting public health. The Arizona Department of Environmental Quality (ADEQ) sponsored a study to determine possible locations for placing monitors to measure DPM (Hafner et al., 2005). Suitability modeling was used to predict areas of high DPM emissions within Phoenix, Arizona. Maps of emission sources, emissions activity data, and meteorology were combined within a GIS model to produce a composite map identifying regions where DPM emissions are likely to be high.

APPENDIX B

FURTHER ANALYSIS EXAMPLES

Examples of three of the more complex analysis techniques (Principal Component Analysis, Emission Inventory, and Suitability Modeling) discussed in Section 3 of the Guidance Document are incorporated in this appendix.

B.1 PRINCIPAL COMPONENT ANALYSIS

Factor analysis, the most common of which is principal component analysis (PCA), is a statistical method to determine the underlying relationships among a group of variables. In the air quality field, the typical application of PCA is on a matrix of *i* ambient samples with concentrations of *j* species, with the goal of resolving a number of factors that describe the relationships among the species over the course of the samples, i.e., what species have similar sources or chemical mechanisms. However, this approach can also be applied to resolve relationships among sites that collect the same parameter such as $PM_{2.5}$. In this case, PCA is conducted on a matrix of *i* ambient sample dates by *l* sites, and factors of sites that have similar temporal variation are determined.

With a valid data set of concentrations of a parameter (i.e., $PM_{2.5}$) in *i* ambient sample dates for *l* sites, or for a set of *j* species in *i* ambient samples, these data can be brought into a statistical package such as SAS, SYSTAT, etc. PCA is a standard statistical application in nearly all statistical packages. Typically, PCA has two options: use of a correlation or a covariation measure. The correlation measure is more often used with ambient data sets because it is not influenced by differences in the magnitude of concentration levels. For a single parameter over multiple sites, this is less of a concern than when there are many species over a wide range of concentration levels (e.g., sulfate vs. trace levels of chromium), but correlation should still be used first. Additionally, there is the option for rotation in the analysis, which allows more realistic factors to be determined. Varimax rotation is most commonly used.

Following is an example of a PCA application that was performed for the Central Regional Air Planning Association (CENRAP) as part of a larger project to identify the causes of regional haze in Class I areas in the CENRAP states.

SPATIOTEMPORAL ANALYSIS

INTRODUCTION

The objective of this task is to identify subregions within CENRAP where aerosol extinction and concentrations of $PM_{2.5}$ components significantly covary in space and time. This analysis will help in selecting representative sites for further analysis which will eliminate the need to model and characterize every site. This task uses recent speciated $PM_{2.5}$ data for 2002-2003 collected as part of the Interagency Monitoring of Protected Visual Environments (IMPROVE) program. The primary tool used in this task is principal component analysis (PCA), with varimax rotation. PCA was applied to identify groups of sites that have similar variance of aerosol extinction (by b_{ext}) or a given species concentration (e.g., organic carbon [OC], nitrate, sulfate, etc.) using data from all sites (Lehman et al., 2004; Eder et al., 1993). The analyses performed in this task built on previous work conducted in Phase I by Desert Research Institute, in which areas of covariance of $PM_{2.5}$ concentrations in the CENRAP and Western Regional Air Partnership (WRAP) regions were identified. The results of this task are sets of sites (i.e., subregions of CENRAP) that share characteristically varying air quality on the 20%-worst and 20%-best visibility days. Representative sites for each subregion are also selected for detailed analyses in later tasks.

METHOD

One-in-three-day IMPROVE data for 2002-2003 at 23 sites in the CENRAP region were obtained from the IMPROVE web site. Basic quality control (QC) was conducted by comparing the measured $PM_{2.5}$ mass to the reconstructed fine mass (RCFM) for every sample at every site (Hafner, 2003). If the comparison showed that the measured mass and RCFM were not within 50%-150%, that sample was flagged as suspect and not used in subsequent data analyses. Forty-four samples were flagged as suspect from this check. Next, the 20%-worst and 20%-best visibility days at each site for 2002-2003 were determined from visibility extinction (b_{ext}). All days when at least one site experienced a 20%-worst visibility day were combined in one subset, and all days when at least one site experienced a 20%-best visibility day were combined in another subset. PCA analysis was then conducted using varimax rotation, which is typically used in such analyses, for the 20% of days with the largest aerosol extinction value and the 20% of days with the smallest aerosol extinction value. This analysis was repeated for sulfate, OC, and nitrate concentrations.

PCA RESULTS FOR AEROSOL EXTINCTION

Results are given in **Table 1** and **Figures 1 and 2**. Six and five subregions were identified from the aerosol extinction on the 20%-worst and 20%-best visibility days, respectively:

- An Upper Midwest subregion, consisting of sites in southern Iowa, Missouri, and eastern Kansas.

- The Western Plains, which includes Big Bend National Park but not the Guadalupe Mountains.
- The Guadalupe Mountains, which consistently show only a loose relationship with Big Bend National Park and other CENRAP sites.
- Minnesota, consisting of the border sites Voyageurs and Boundary Waters.
- Southeastern Plains, which includes sites in Louisiana and southern Arkansas.
- On the 20%-worst visibility days, the additional subregion was a "transition zone" between the western plains, upper Midwest and Southeastern Plains and consisted of Upper Buffalo and Wichita Mountains.

From these results, four representative sites were selected: Cedar Bluff, Kansas, for the Western Plains; Sikes, Louisiana, for Southeastern Plains; Hercules-Glades, Missouri, for the Upper Midwest; and Voyageurs, Minnesota, for Minnesota. The transition zone sites are approximated by the representative sites.

The selection of the representative sites was confirmed by comparing the number of 20%-worst and 20%-best visibility days each site had in common with the other sites in its subregion. Minnesota only had two sites, so Voyageurs was selected because it provided more data than did Boundary Waters. In the Upper Midwest, El Dorado Springs was the most representative site, followed by Tallgrass and Hercules-Glades. However, Hercules-Glades was selected because it provided twice as much data as did El Dorado Springs and is still very representative of the region. In the Western Plains, all sites but Big Bend shared nearly all the same days, and Cedar Bluff was the most representative. In the Southeastern Plains, Sikes was the most representative site in its subregion.

PCA RESULTS FOR $PM_{2.5}$ COMPONENTS

In addition to aerosol extinction, groupings among sites for dominant aerosol components were explored with PCA. This helps us understand the underlying variability of the PCA analysis on aerosol extinction, how representative the selected sites are, and the extent of regional versus local effects.

PCA results using OC, NO_3 and SO_4 on the 20%-worst and 20%-best visibility days are shown in **Figures 3 through 8**. Results were consistent with the aerosol extinction analysis, but showed some underlying trends that will be useful in later analyses:

- Nitrate concentrations varied more on a local level than on a regional level; five to seven factors were found for nitrate. The Upper Midwest factor identified by b_{ext} was split into two, which may be due to the greater availability of ammonia for ammonium nitrate formation in Iowa compared to Missouri.
- Sulfate showed a distinct regional character, with the Minnesota, Upper Midwest, Transition Zone and Southeastern Plains grouped together. The Western Plains, Big Bend, and Guadalupe visibility trends are likely distinguished from the other sites by sulfate differences.

- OC showed results similar to aerosol extinction results, except that the Western Plains and Minnesota were grouped together. This may be indicative of a "western" OC influence in these subregions versus a more localized OC influence in the eastern subregions.

REFERENCES

Eder B.K., Davis J.M., and Bloomfield P. (1993) A characterization of the spatiotemporal variability of non-urban ozone concentrations over the eastern United States. *Atmos. Environ.* **27A**, 2645-2668.

Hafner H.R. (2003) 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.

Lehman J., Swinton K., Bortnick S., Hamilton C., Baldridge E., Eder B., and Cox B. (2004) Spatio-temporal characterization of tropospheric ozone across the eastern United States. *Atmos. Environ.* **38**, 4357-4369.

Table 1. PCA results on 20%-worst and 20%-best visibility days for aerosol extinction.

Subregion	% Variance on Worst-20% Days	% Variance on Best-20% Days	Representative Site
Minnesota	12	8	Voyageurs
Upper Midwest	36	42	Hercules-Glades
Western Plains	16	23	Cedar Bluff
Transition Zone	11	–	–
Southeastern Plains	10	12	Sikes
Guadalupe Mountains	7	9	–

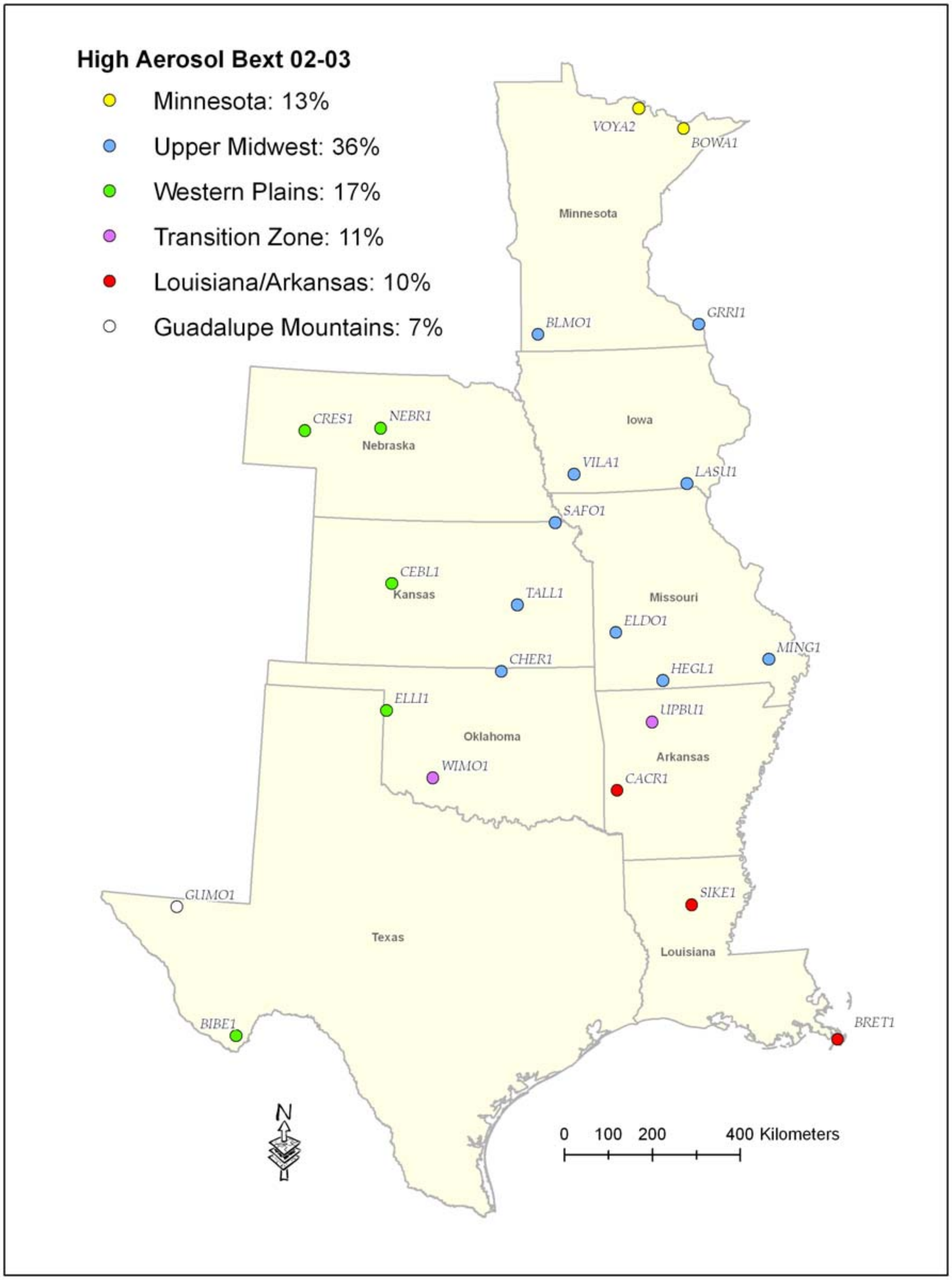


Figure 1. PCA results for aerosol extinction on worst 20% visibility days 2002-2003. The colored dots represent the groups determined by PCA. The percentage of data variability accounted for by each group is also shown.

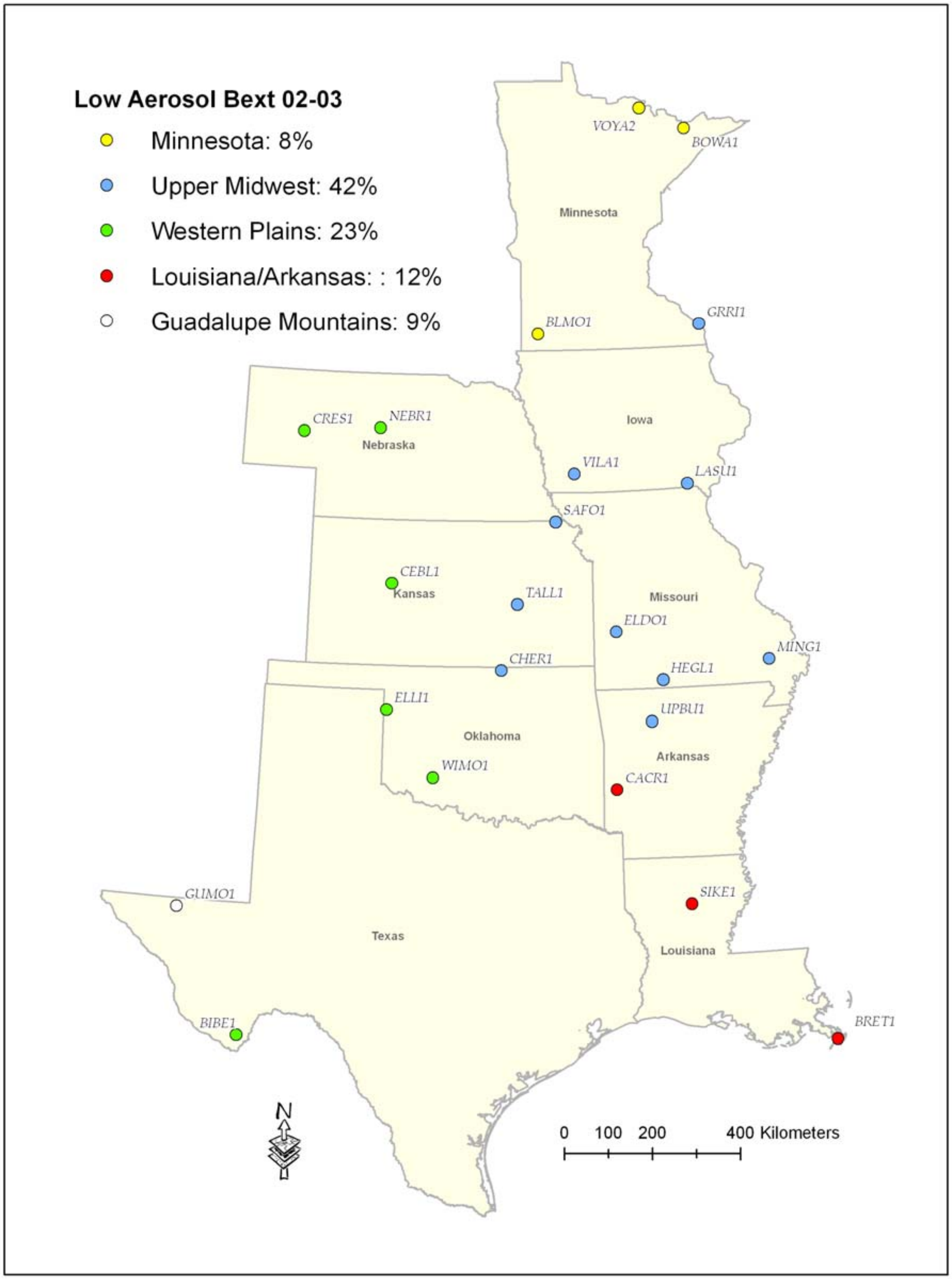


Figure 2. PCA results for aerosol extinction on best 20% visibility days 2002-2003. The colored dots represent the groups determined by PCA. The percentage of data variability accounted for by each group is also shown.

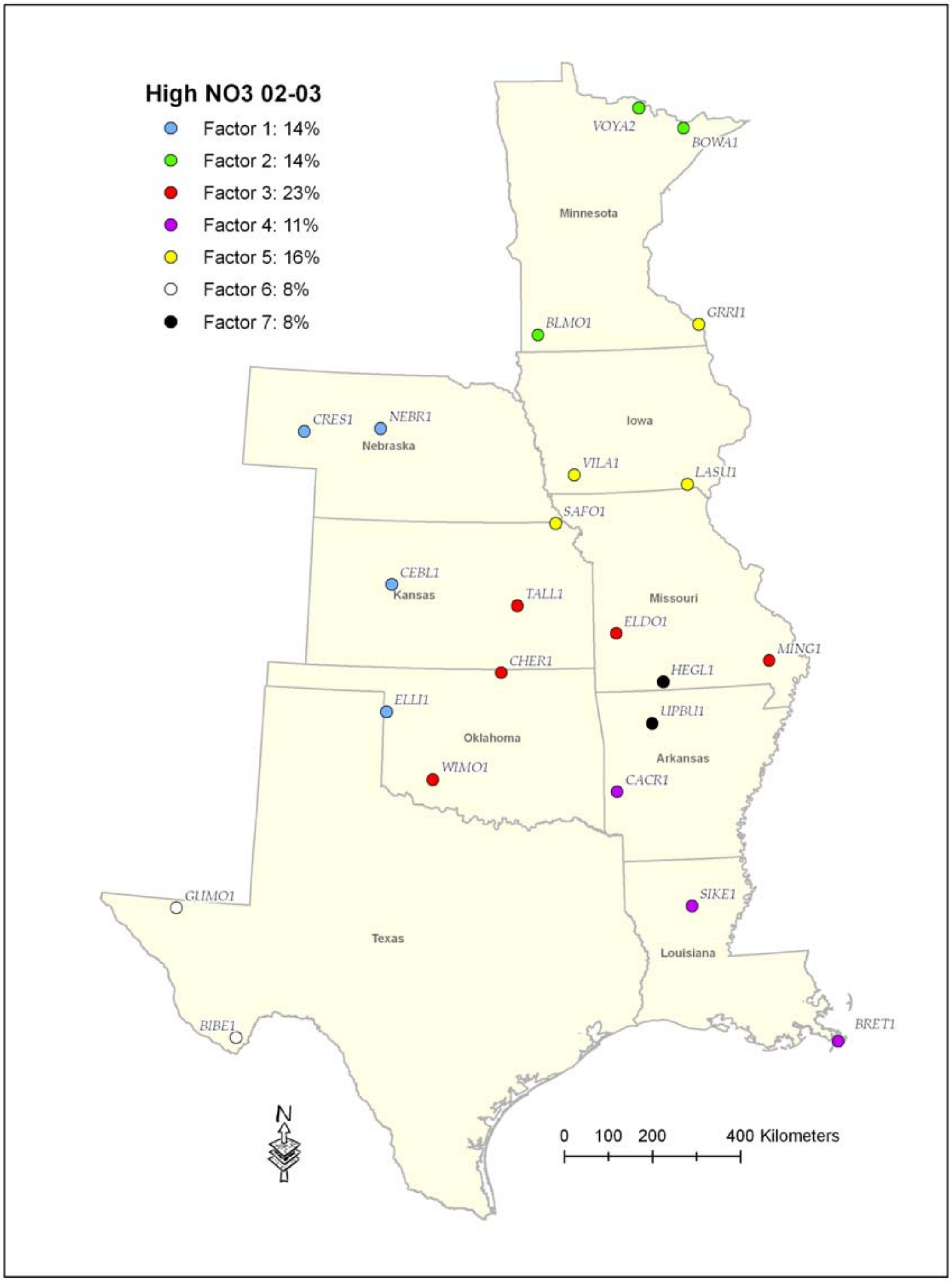


Figure 3. PCA results for nitrate on worst 20% visibility days 2002-2003. The colored dots represent the groups determined by PCA. The percentage of data variability accounted for by each group is also shown.

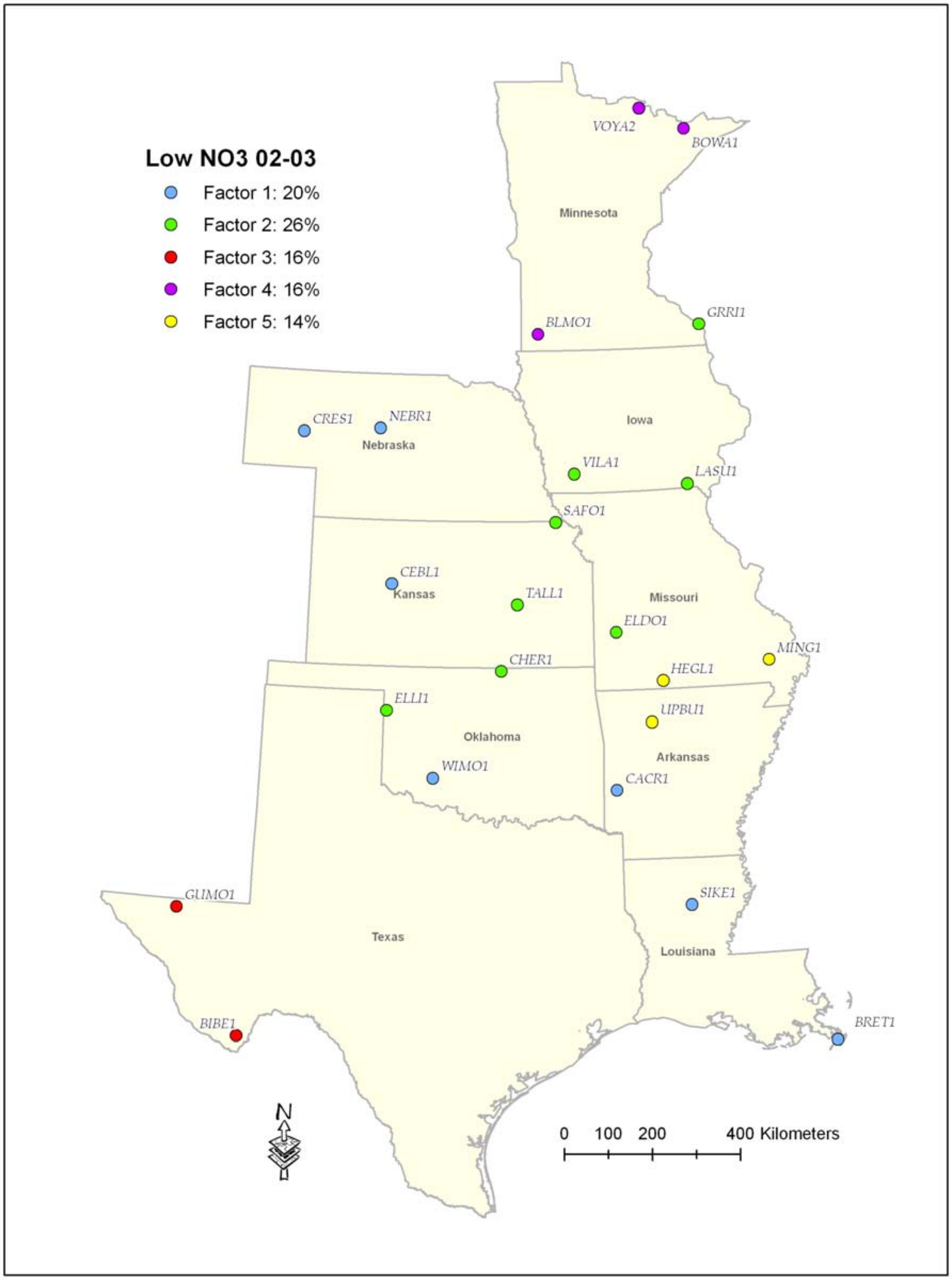


Figure 4. PCA results PCA results for aerosol extinction on best 20% visibility days 2002-2003. The colored dots represent the groups determined by PCA. The percentage of data variability accounted for by each group is also shown.

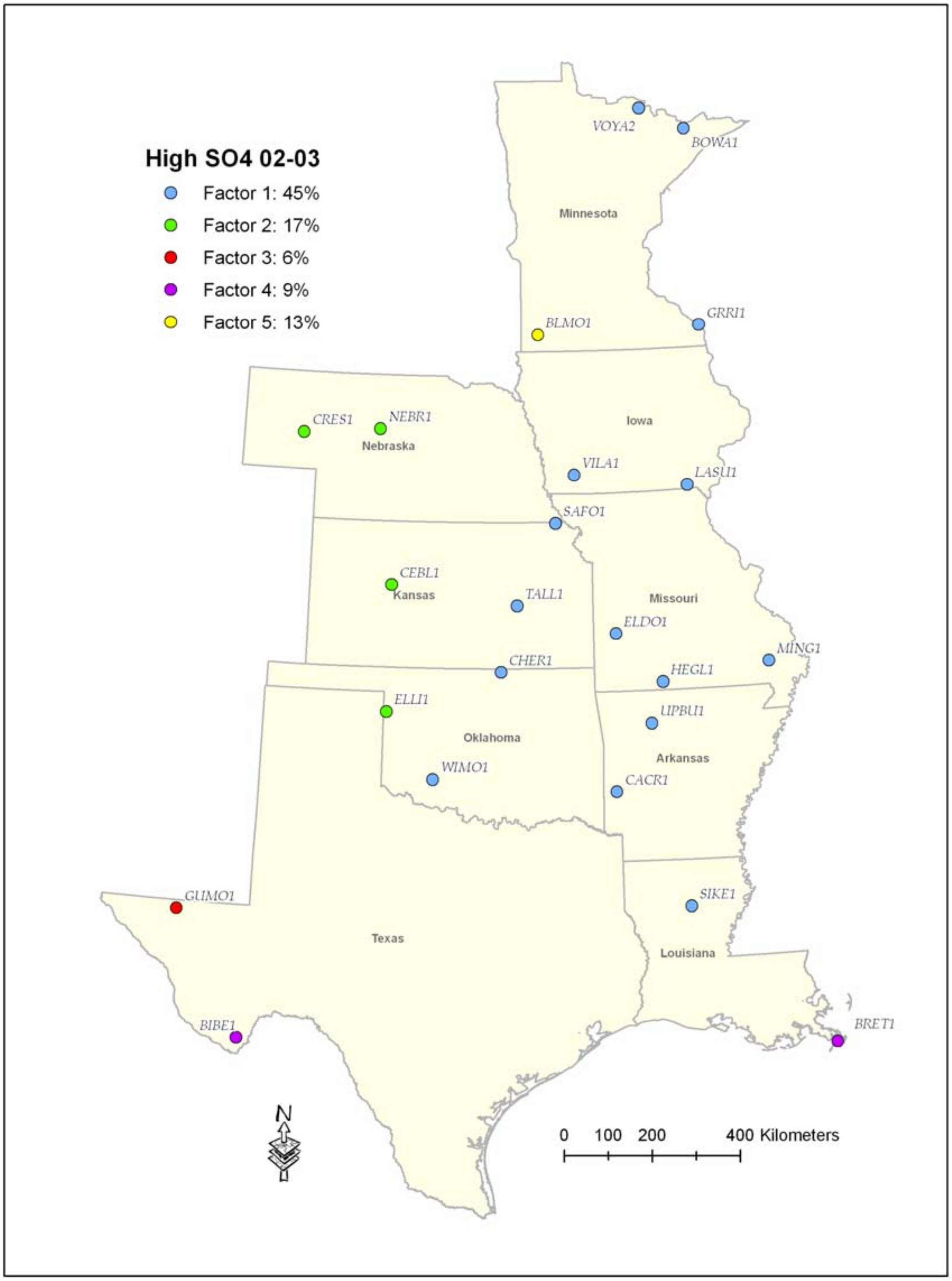


Figure 5. PCA results PCA results for sulfate on worst 20% visibility days 2002-2003. The colored dots represent the groups determined by PCA. The percentage of data variability accounted for by each group is also shown.

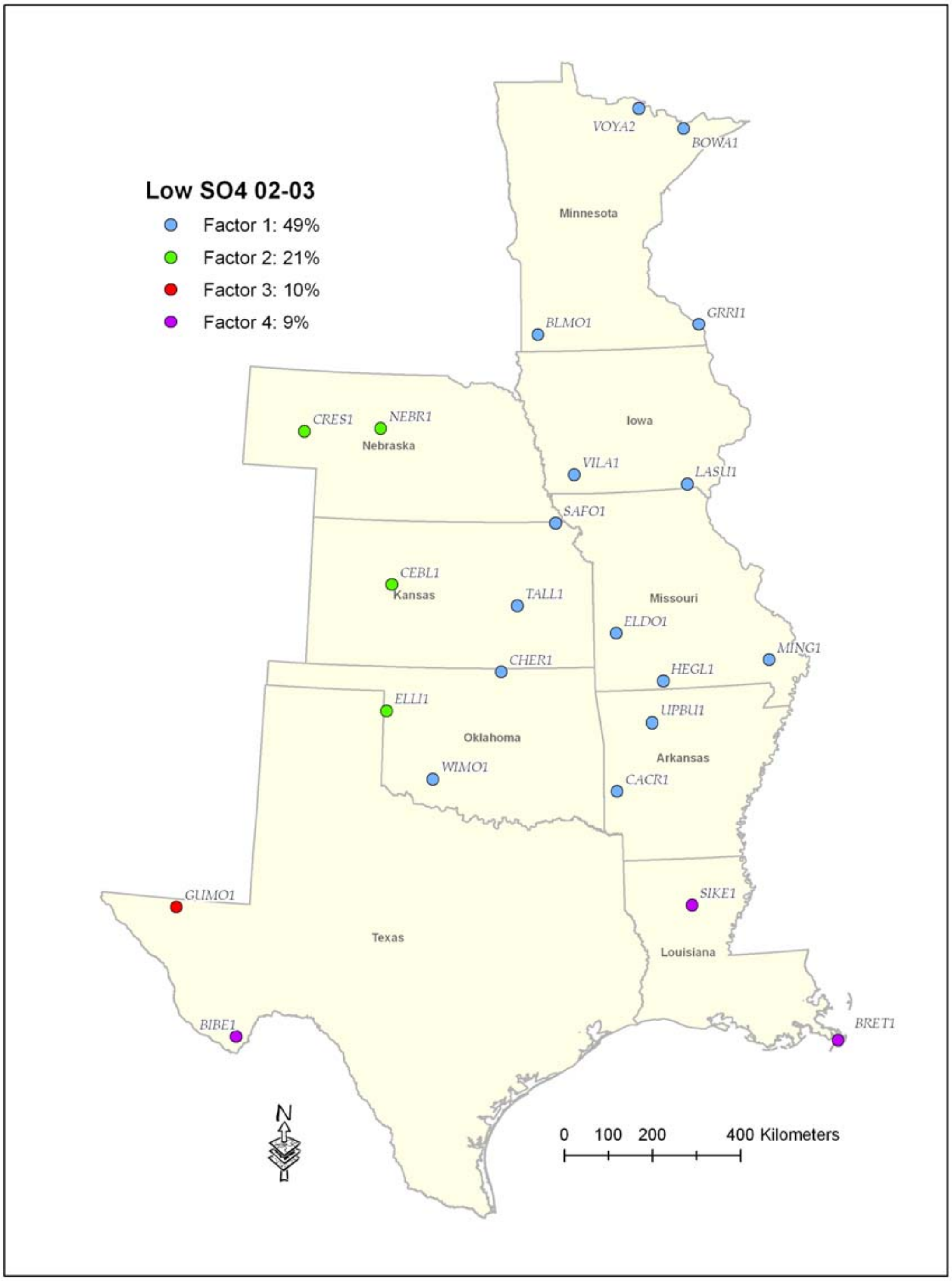


Figure 6. PCA results PCA results for sulfate on best 20% visibility days 2002-2003. The colored dots represent the groups determined by PCA. The percentage of data variability accounted for by each group is also shown.

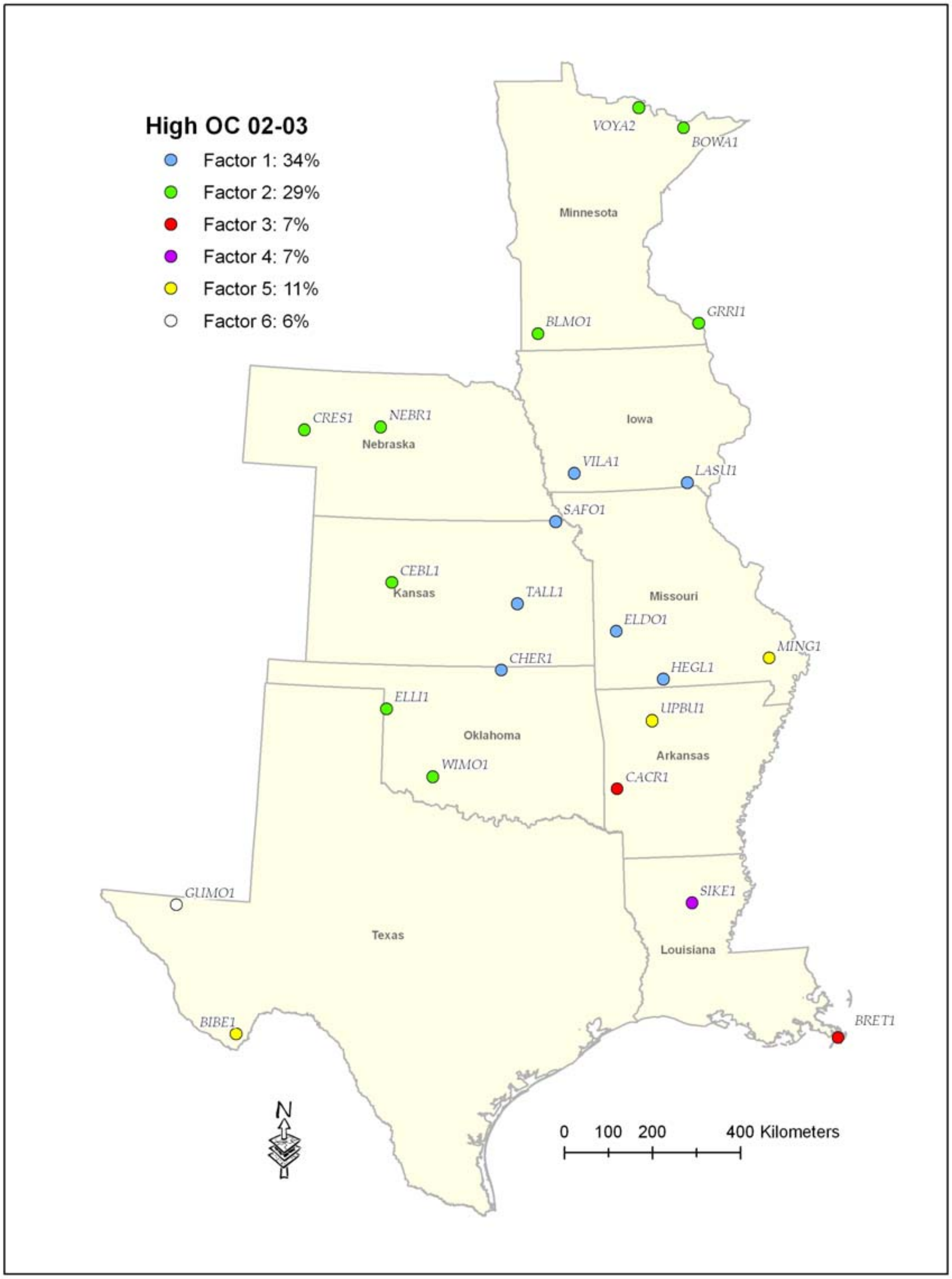


Figure 7. PCA results PCA results for OC on worst 20% visibility days 2002-2003. The colored dots represent the groups determined by PCA. The percentage of data variability accounted for by each group is also shown.

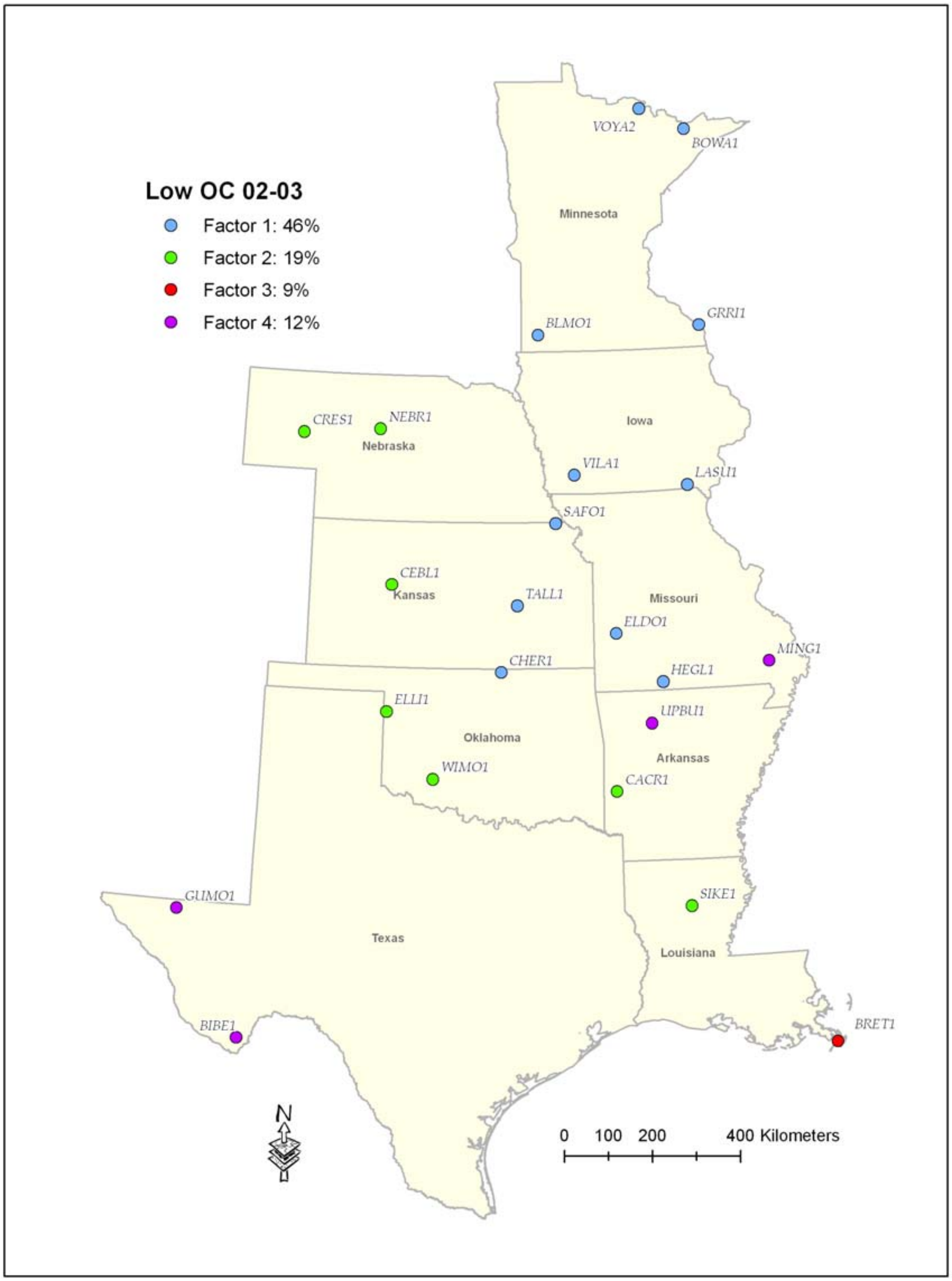


Figure 8. PCA results PCA results for OC on best 20% visibility days 2002-2003. The colored dots represent the groups determined by PCA. The percentage of data variability accounted for by each group is also shown.

B.2 EMISSION INVENTORY

The EPA publishes a National Emission Inventory (NEI) every three years. The NEI resolves criteria air pollutants and hazardous air pollutants to annual county-wide emissions. These inventories are useful for identifying emissions patterns at a national or regional scale; however, for network assessment purposes, it is useful to have an emission inventory at a finer spatial resolution (e.g., a few kilometers per grid cell). Emission inventories at sub-county resolution are not readily available. The attached report is an example of a project to create a spatially resolved toxics emission inventory in the San Francisco Bay Area. In addition to allocating county-wide emissions to much smaller grid cells, $PM_{2.5}$ and VOC inventories were speciated into individual toxic species. These higher resolution data enable analyses such as those shown on the Emission Inventory summary sheet in Section 3.

Reference

U.S. Environmental Protection Agency (2005) Technology Transfer Network—Clearinghouse for inventories and emission factors, National Emissions Inventories for the U.S. Available on the Internet at <http://www.epa.gov/ttn/chief/net/index.html>.

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

PREPARATION OF EMISSION INVENTORIES OF TOXIC AIR CONTAMINANTS FOR THE BAY AREA

Final Report
STI-905019-2771-FR2

By:
Dana C. Sullivan
Stephen B. Reid
Bryan M. Penfold
Elizabeth M. Cozzo
Sonoma Technology, Inc.
1360 Redwood Way, Suite C
Petaluma, CA 94954-1169

Prepared for:
Bay Area Air Quality Management District
939 Ellis Street
San Francisco, CA 94109

August 17, 2005

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF FIGURES

1. INTRODUCTION

The Bay Area Air Quality Management District (BAAQMD) is carrying out the Community Air Risk Evaluation (CARE) program to characterize and reduce health risks from toxic air contaminants (TACs) emitted in the San Francisco Bay Area. In support of the CARE program's goals, screening-level gridded emission inventories of TACs were prepared (including diesel exhaust). These screening-level inventories were assembled top-down from readily available information and represent an early step from which subsequent CARE program activities may be carried forward. They are intended to aid in the identification and prioritization of further inventory development activities, to support initial exposure modeling runs, and to facilitate selection of a study community in the Bay Area. Although the screening-level inventories represent many of the TACs and emissions sources that are important in the Bay Area, they should be considered "working versions under development"—i.e., inventories that are useful, but should be augmented and improved in consideration of reviewers' feedback or as additional inventory data are developed bottom-up through CARE program activities.

Mass-based and risk-weighted emission inventories of TACs were prepared by acquiring and processing existing and available information sources. These information sources included (1) a county-level area and non-road mobile source inventory of year-2000 emissions of total organic gases (TOG) and particulate matter less than 10 microns (PM_{10}); (2) gridded on-road mobile source inventories of year-2000 summer and winter weekday TOG and PM_{10} emissions in Modeling Emissions Data System (MEDS) format; (3) a year-2000 inventory of TAC emissions from point sources in the BAAQMD; (4) chemical speciation profiles published by the U.S. Environmental Protection Agency (EPA), California Air Resources Board (ARB), and Desert Research Institute (DRI); (5) geographic information systems (GIS) databases acquired from the Association of Bay Area Governments (ABAG) and other sources; (6) ratios of particulate matter less than 2.5 microns ($PM_{2.5}$):PM_{10} calculated from information published by the ARB; and (7) inhalation unit risk factors and reference concentrations available from the ARB and EPA. The resultant inventories should be considered suitable for initial exposure modeling and data analysis with a view toward prioritizing TACs and emissions source categories of concern for future research efforts. Conclusions and recommendations for further improvements to the inventories are presented in Section 5.

2. SUMMARY OF EXISTING INVENTORIES

The BAAQMD's pre-existing inventories of TOG, PM_{10}, and TACs were used as the basis for compiling screening-level TAC emission inventories for the CARE program. **Figure 2-1** illustrates the distribution of total TOG and PM_{10} emissions by major source category (point, on-road mobile, and area/non-road). Section 2 discusses each of the major source categories in greater detail and outlines the techniques used to geographically allocate emissions to the BAAQMD's 2-km × 2-km modeling grid.

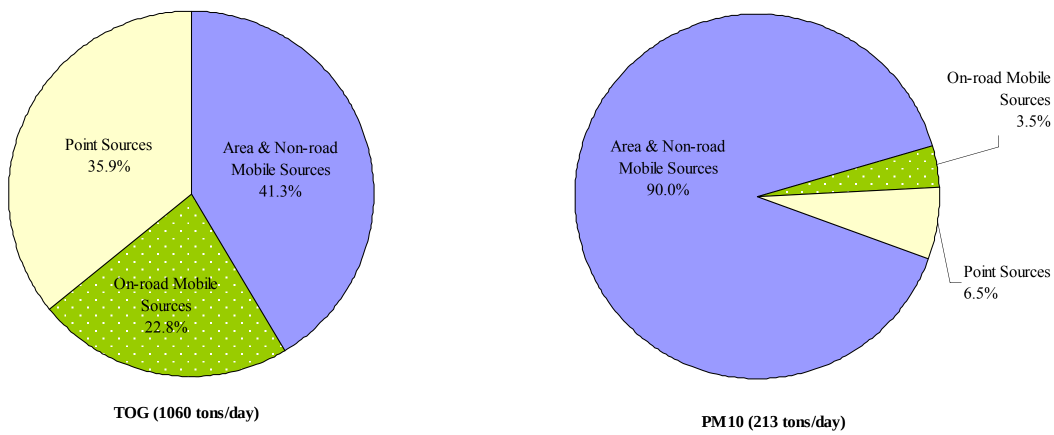


Figure 2-1. 2000 TOG and PM_{10} emissions by major source type for the BAAQMD.

2.1 EMISSIONS FROM AREA AND NON-ROAD MOBILE SOURCES

The BAAQMD provided Sonoma Technology, Inc. (STI) with county-level area and non-road mobile source inventories of year-2000 TOG and PM_{10} emissions for 345 emission inventory codes (EICs). Emission estimates for $PM_{2.5}$ were developed by applying $PM_{2.5}$-to-PM_{10} size fractions recommended by the ARB for each EIC code (see Appendix A). **Figures 2-2 and 2-3** summarize area and non-road mobile source TOG, PM_{10}, and $PM_{2.5}$ emissions by detailed source categories, and **Figure 2-4** shows an emission density plot of area and non-road mobile source TOG emissions. Important sources of TOG include evaporative sources (e.g., petroleum marketing and consumer products) and combustion sources (e.g., off-road equipment and residential fuel combustion). Sources of TOG associated with the decomposition of organic matter (e.g., livestock waste and landfills) emit most TOG in the form of methane—a non-toxic gas. Important sources of PM include combustion sources (e.g., residential fuel combustion and off-road equipment). Sources of fugitive dust emit most PM in the form of non-toxic geologic material.

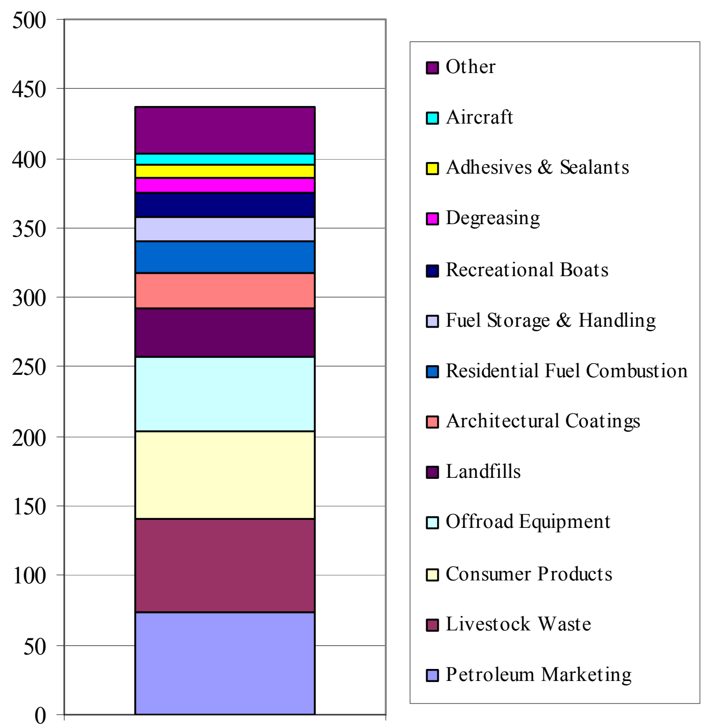


Figure 2-2. Area and non-road mobile source TOG emissions (tons/day) by source category.

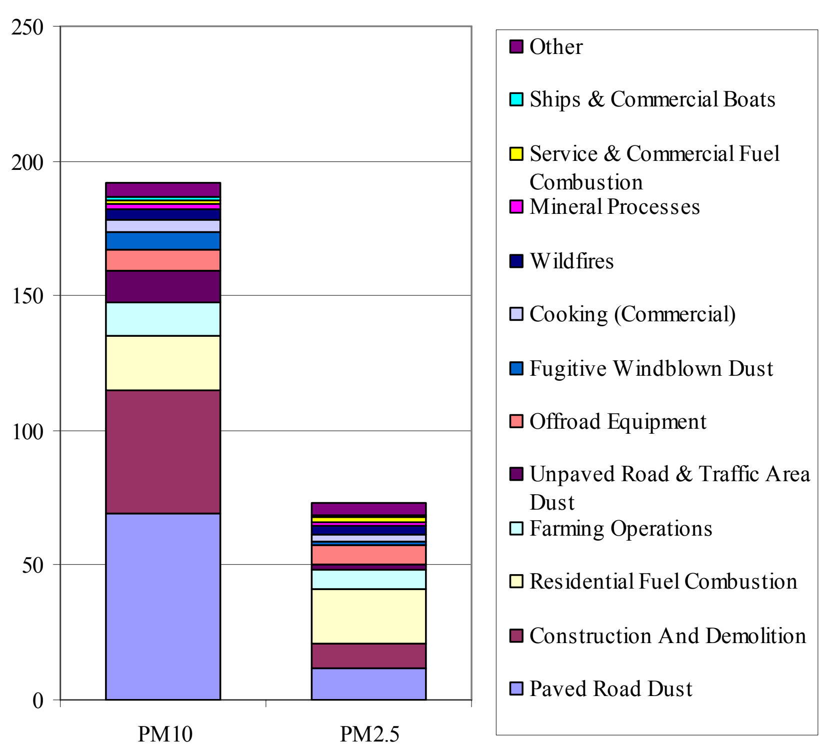


Figure 2-3. Area and non-road mobile source PM emissions (tons/day) by source category.

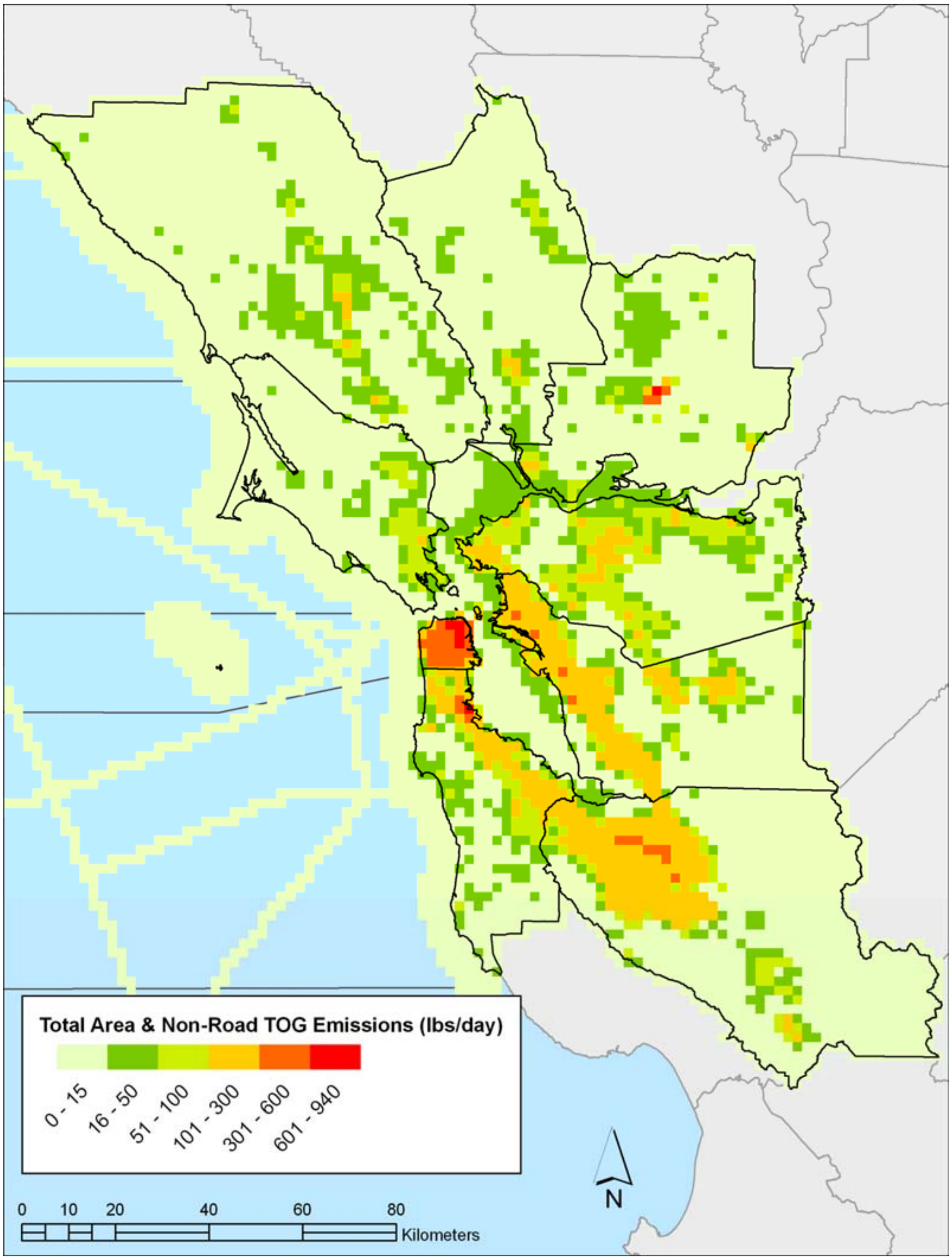


Figure 2-4. Emission density plot of area and non-road mobile source TOG emissions.

County-level area and non-road mobile emissions were geographically distributed using GIS databases. GIS databases with suitable spatial resolutions were selected as surrogates for the locations of emissions sources. County-wide emissions were allocated to individual grid cells proportionally according to the spatial patterns of the surrogate GIS data. Examples of some GIS databases employed for the process include land use area (e.g., industrial land use), line length (e.g., railroad track length), line density (e.g., roadway traffic activity), and point count (e.g., number of dry cleaning locations). Allocation factors were developed for individual grid cells by processing GIS data within a customized ArcGIS Visual Basic (VBA) program that outputs allocation factors by grid cell to Microsoft Access database tables. Four basic types of spatial allocation calculations were used to develop the spatial surrogates applied to the inventory, based on the type of GIS data (i.e., polygon, line, or point).

A variety of GIS data sets were used to geographically distribute county-wide emissions. The most often-used data set for this process was the Existing Land Use – 2000, developed by the ABAG. The ABAG land use database incorporates the U.S. Geological Survey National Land Cover Dataset (NLCD) as well as county assessors' data on land use. Another important data source was the Central California Ozone Study (CCOS) gridded surrogates project, (Funk et al., 2001), which included representations of the geographic locations of businesses derived from records of business addresses. Addresses for auto body shops, dry cleaners, and variety of other types of businesses were geocoded to estimate geographic locations, which Funk et al. (2001) used in turn to calculate gridded spatial allocation factors. In total, 46 surrogates were developed for spatially allocating the area and non-road sources. The details about GIS data sources and methods used for spatial allocation are provided in Appendix B.

2.2 EMISSIONS FROM ONROAD MOBILE SOURCES

The BAAQMD provided STI with gridded on-road mobile source inventories of TOG and PM_{10} emissions in MEDS format. BAAQMD generated summer and winter MEDS files for 2000, and these two inventories were averaged by grid cell to produce an annualized on-road emissions inventory. As with emissions from area and off-road mobile sources, emission estimates for $PM_{2.5}$ were developed from the on-road PM_{10} inventory by applying ARB $PM_{2.5}$-to-PM_{10} size fractions recommended by the ARB. **Figures 2-5 and 2-6** summarize on-road mobile source TOG, PM_{10}, and $PM_{2.5}$ emissions by source category, and **Figure 2-7** shows an emission density plot of the gridded on-road mobile source TOG inventory provided by BAAQMD. Evaporative losses and vehicle exhaust emissions are important sources of TOG and PM.

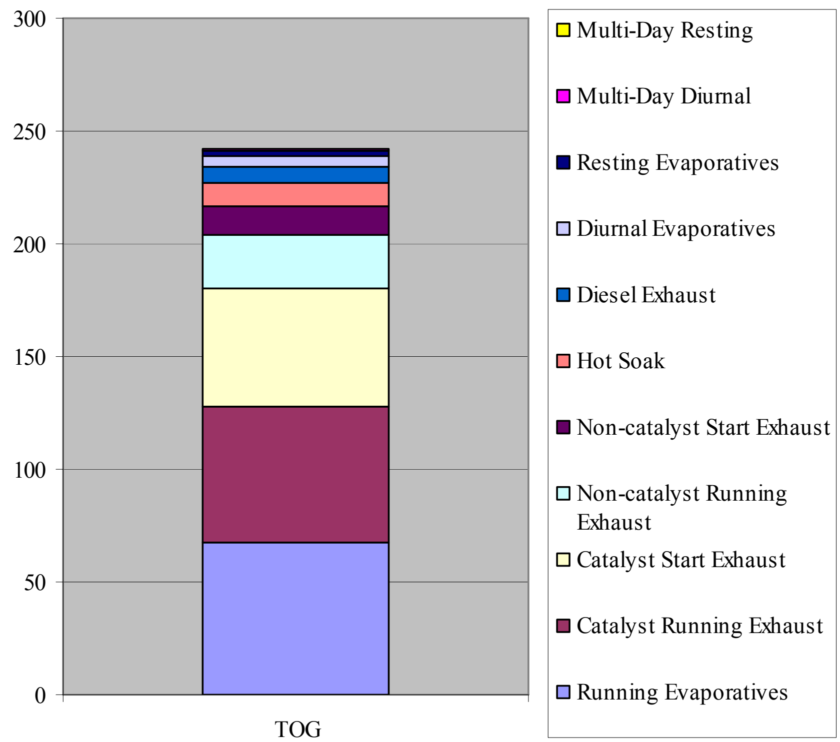


Figure 2-5. On-road mobile source TOG emissions (tons/day) by source category.

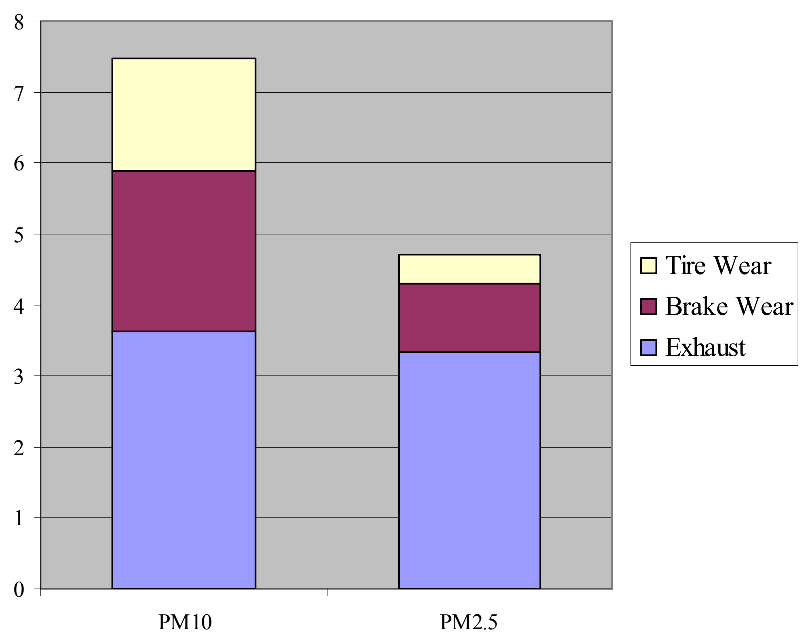


Figure 2-6. On-road mobile source PM emissions (tons/day) by source category.

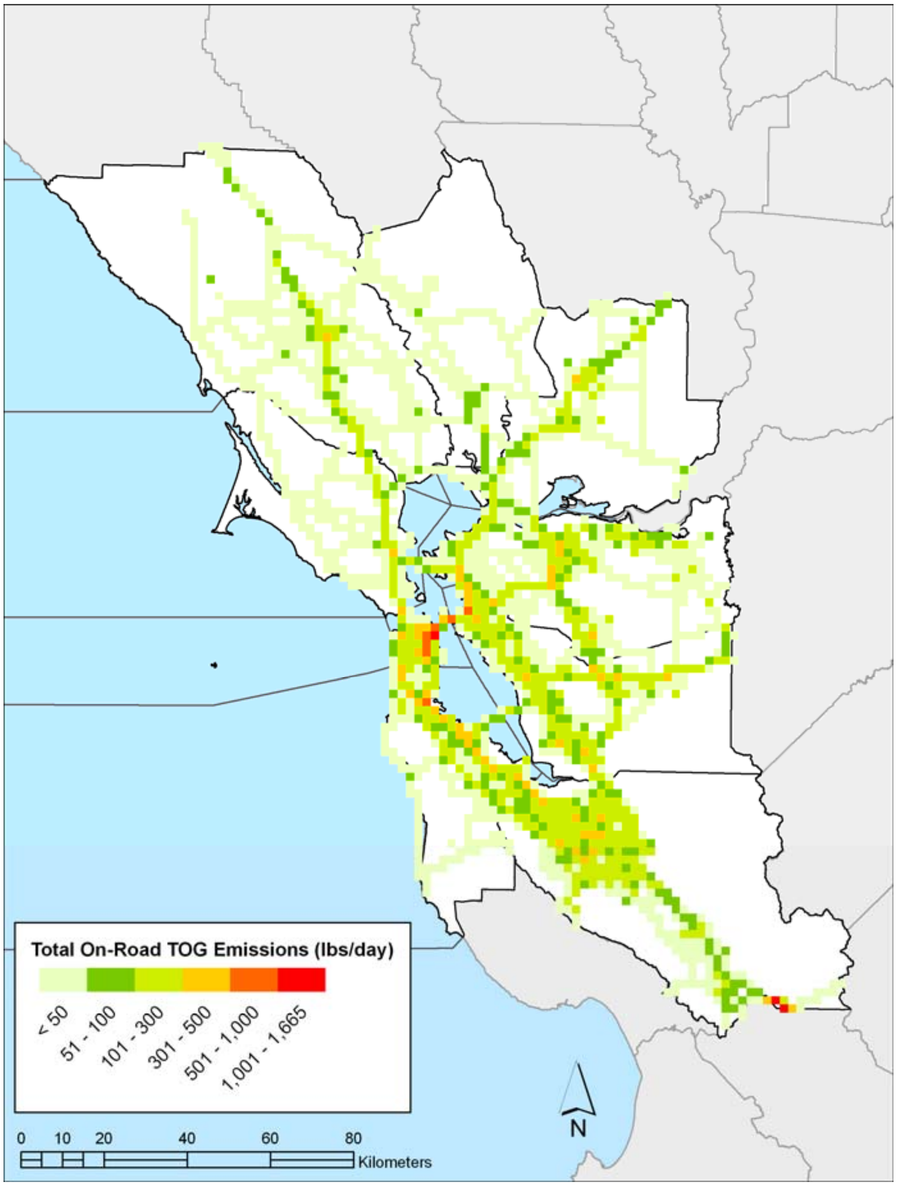


Figure 2-7. Emission density plot of on-road mobile source TOG emissions.

2.3 EMISSIONS FROM POINT SOURCES

For the purposes of direct comparison to the data summarized in Sections 2.1 and 2.2, STI downloaded year-2000 TOG, reactive organic gases (ROG) and PM_{10} point-source emissions from ARB's Facility Search Engine. **Figures 2-8 and 2-9** summarize point source ROG, PM_{10}, and $PM_{2.5}$ emissions by facility. (ROG emissions are displayed because the TOG inventory is dominated by landfills, which emit most TOG in the form of methane—a non-toxic gas). Petroleum refineries are important sources of ROG and PM_{10}, and electrical generation facilities and landfills are also important sources of PM_{10}.

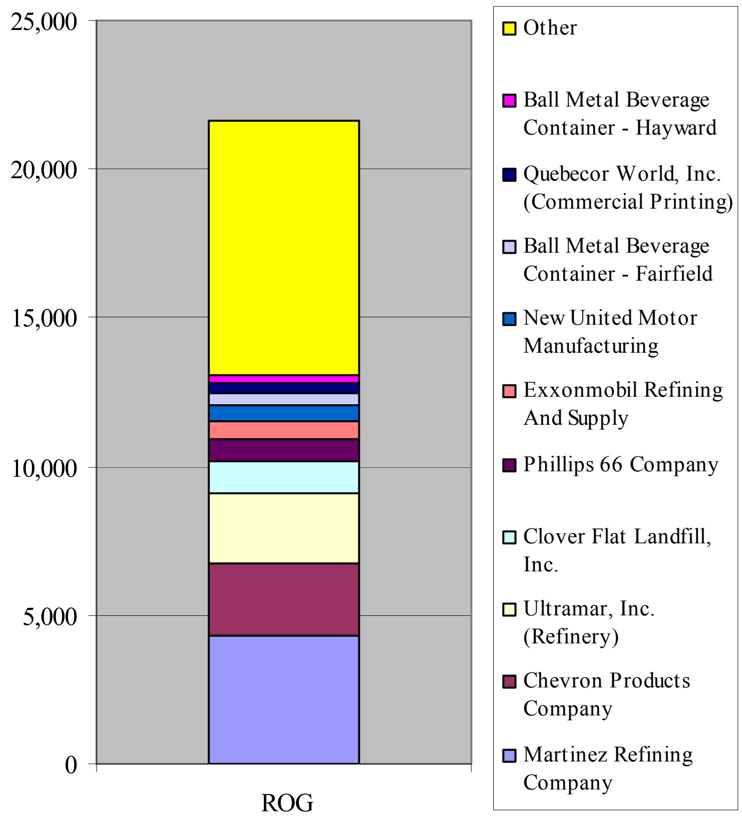


Figure 2-8. Point source ROG emissions (tons/year) by facility.

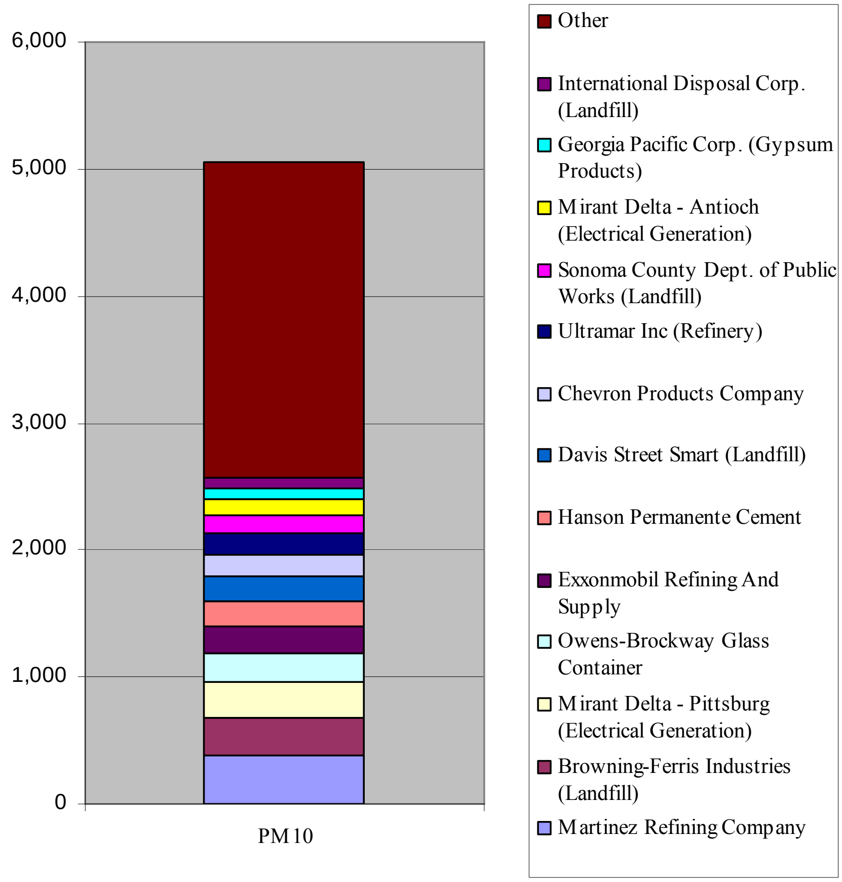


Figure 2-9. Point source PM_{10} emissions (tons/year) by facility.

In order to assign a geographical location to individual point sources, facility addresses reported in the BAAQMD's point source inventory were geocoded. The geocoding was achieved using the Tele Atlas EZ-Locate software. The Tele Atlas' geocoding process relies on a combination of postal standardization software, positionally accurate maps, and advanced geocoding techniques to generate the most accurate geocode coordinates available. Each point source in the BAAQMD's inventory file was assigned a geographic location and a geocode match type field. The match type is determined according to the accuracy of the geocoded address. A match type of "1" is the best match type possible, where the geographic location is accurate to an exact house number within a single side of a single street block. A match type of "4" is accurate only to the 5-digit zip code centroid. Of 3,359 unique addresses in the BAAQMD's point source inventory file, 4% (137 addresses) received a match type "4" and 96% (3,222 addresses) received a match type "1". The estimated locations of point sources with a match type "4" are inexact and should be corrected or taken into consideration during exposure modeling and risk assessment efforts. **Figure 2-10** illustrates the geographic distribution of benzene emissions from point sources. Details about the point source inventory geocoding match type results are documented in Appendix B.

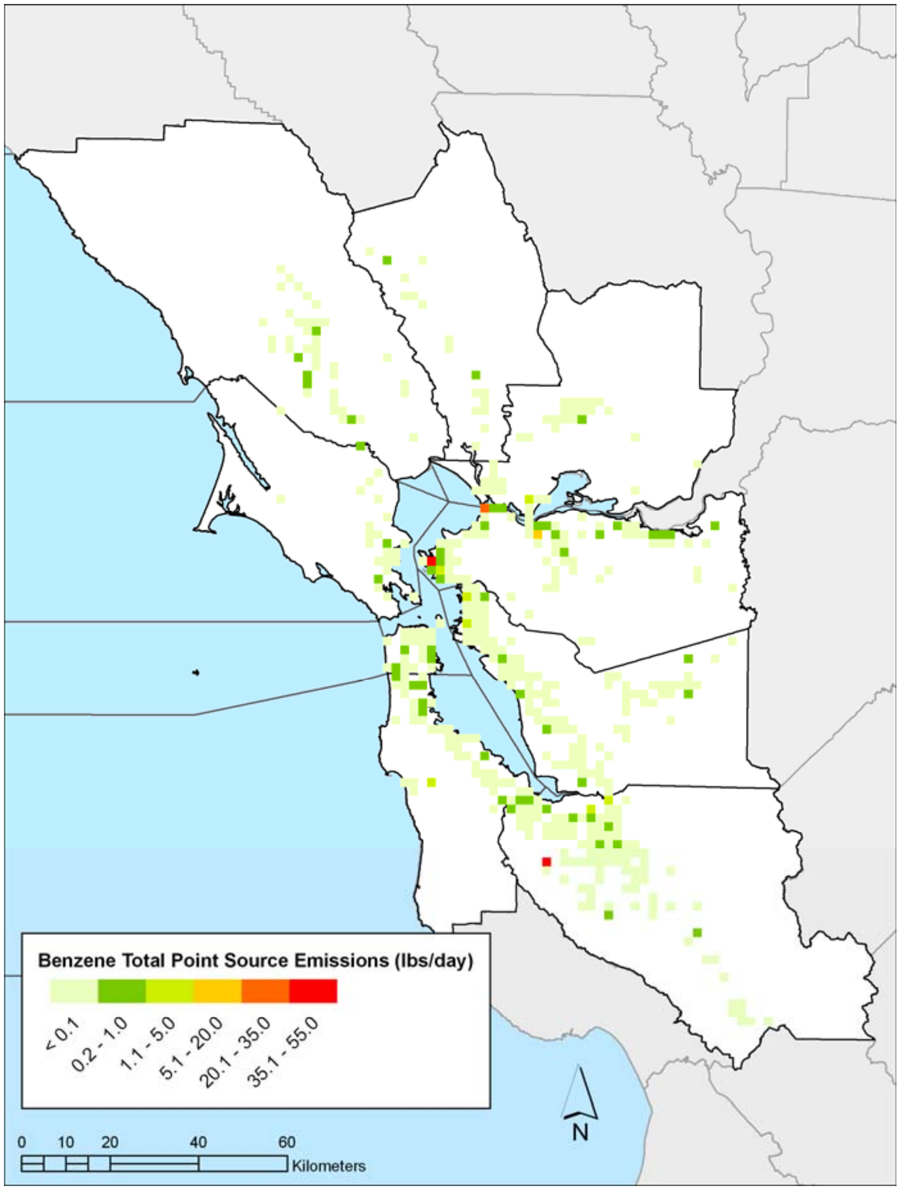


Figure 2-10. Emission density plot of benzene emissions from point sources.

3. DEVELOPMENT OF TOXIC AIR CONTAMINANT INVENTORIES

Several data processing steps were followed to prepare the inventories for chemical speciation. STI assigned speciation profiles to each EIC or source classification code (SCC) in the area, non-road mobile, and on-road mobile source TOG and PM inventories. The ARB has developed a database of TOG and PM speciation profiles, as well as a cross-reference file that indicates which TOG or PM profile is assigned to each EIC code. In most cases, speciation profile-to-EIC code assignments recommended by the ARB were followed. However, where no speciation profile (or a composite profile) was recommended for a given source category by the ARB, appropriate profiles from the DRI or the EPA's Speciate 3.2 database were utilized. (Appendix B contains a listing of the speciation profiles used in this study and a cross-reference table matching these profiles to each EIC/SCC code in the inventories). The BAAQMD provided STI with a year-2000 inventory of TACs from point sources in the Bay Area; therefore, no processing steps were required to accomplish chemical speciation of the point source inventories.

The selected speciation profiles were used to transform TOG and PM emissions into individual chemical species so that TAC emissions from area, non-road mobile, and on-road mobile sources could be estimated. These estimates were combined with the inventory of TAC emissions from point sources provided by BAAQMD to form a complete inventory of TACs. TAC emissions by pollutant and source category can be seen in **Figures 3-1 and 3-2**. Note that Figures 3-1 and 3-2 illustrate emissions on a mass basis; however, the toxicity and the geographic distribution of each TAC must be considered in order to evaluate its potential for posing human health risks. Section 4 discusses measures of toxicity and illustrates the geographic distributions of several potentially important TACs.

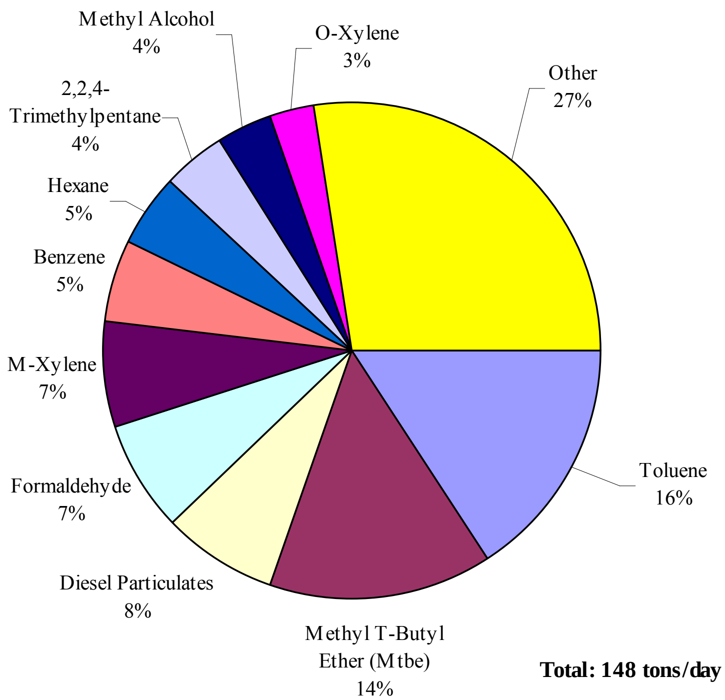


Figure 3-1. Average daily TAC emissions by species.

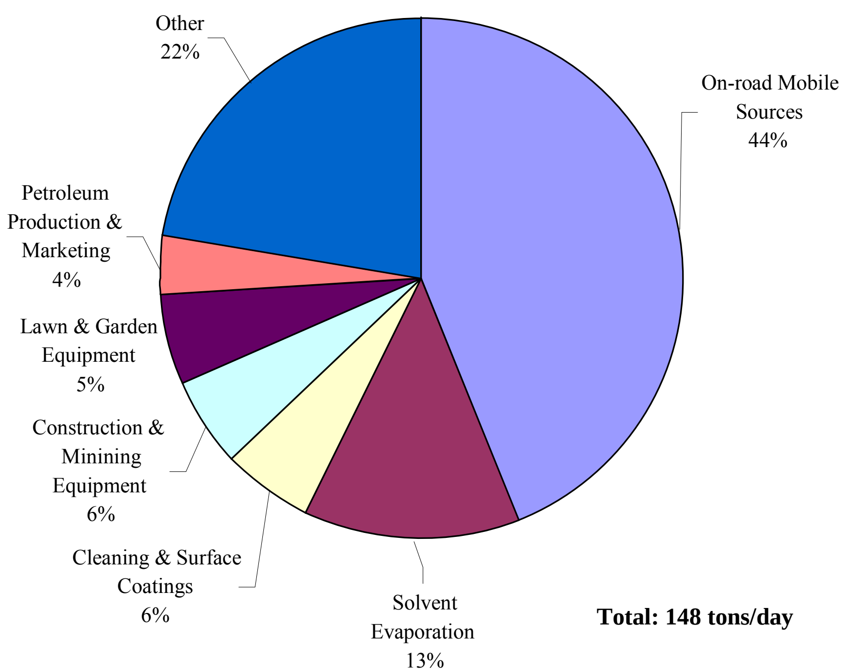


Figure 3-2. Average daily TAC emissions by detailed source category.

4. DEVELOPMENT OF RISK-WEIGHTED TOXIC INVENTORIES

To develop risk-weighted emission inventories, STI applied available cancer unit risk (UR) factors and non-cancer reference concentrations (RfC) for the inhalation exposure pathway. UR factors estimate the expected change in the rate of observed adverse effects per unit change in dose (or air concentration). A RfC is a regulatory definition that indicates the dose at which no adverse effects are expected plus a safety margin allowing for measurement uncertainty plus another safety margin based on professional toxicologists' judgment. UR factors and RfCs were compiled from the following information sources in declining order of preference: ARB in conjunction with EPA's Office of Environmental Health Hazard Assessment (OEHHA), the EPA's Integrated Risk Information System (IRIS), and the EPA's Technology Transfer Network. Secondary sources were used to estimate factors for important TACs not available in the preferred references. When URs or RfCs were reported as a range (e.g., Benzene, CAS# 71432, IRIS), the high end of the range was used to prepare the risk-weighted emission inventory and range-related uncertainties were calculated and documented. Details about the selection, application, and calculation of quantifiable uncertainties are documented in Appendix C.

Mass-based emissions for all TACs were converted to risk-weighted emissions for cancer, chronic, and acute risks. Risk-weighted emissions are reported in units of "mass equivalents per unit time". For risks of cancer due to inhalation, the mass equivalent of a specific TAC is the estimated mass of hypothetical compound "X" that poses a cancer risk equal to that of the emitted mass of the TAC of interest (where "X" is defined as having a UR factor equal to one case of cancer per hundred thousand individuals exposed per unit change in ambient concentration). Thus, risk-weighted emissions for cancer due to inhalation exposure are calculated according to Equation 4-1.

$$\text{Emissions} \times UR_i \div UR_X = \text{Risk-Weighted Emissions} \qquad \textbf{(4-1)}$$

where:

Emissions = Mass-based emissions of a TAC species *i*; **lbs/day**

UR_i = UR factor for TAC species *i*, or the expected change in the number of cases of cancer per million individuals exposed per unit change in ambient concentration; $\mathbf{10^{-5} \cdot (\mu g/m^3)^{-1}}$

UR_X = UR factor for a hypothetical compound "X" $\equiv 1 \times 10^{-5} \cdot (\mu g/m^3)^{-1}$

Risk-Weighted Emissions = Equivalent emissions of hypothetical compound "X", which would be expected to pose a risk equal to that of the emissions of the TAC species *i*; **equivalent lbs/day**

For non-cancer risks due to inhalation (whether acute or chronic), the mass equivalent of a specific TAC is the estimated mass of hypothetical compound "Y", which would be expected to pose a risk equal to that of the emitted mass of the TAC of interest (where "Y" is defined as having an RfC equal to unity). Thus, risk-weighted emissions for acute or chronic effects due to inhalation exposure are calculated according to Equation 4-2.

$$\text{Emissions} \div RfC_i \times RfC_Y = \text{Risk-Weighted Emissions} \quad \textbf{(4-2)}$$

where:

Emissions = Mass-based emissions of a TAC species i; **lbs/day**

RfC_i = RfC for TAC species i; **μg/m³**

RfC_Y = RfC for hypothetical compound "Y" ≡ 1 μg/m³

Risk-Weighted Emissions = Equivalent emissions of hypothetical compound "Y", which would be expected to pose a risk equal to that of the emissions of the TAC species i; **equivalent lbs/day**

Figures 4-1 through 4-6 show risk-weighted emissions by pollutant and source category for cancer-related, chronic, and acute effects caused by inhalation exposure. Chromium and diesel particulate matter comprise about 90% of cancer risk-weighted emissions. On-road mobile sources and construction-related activities contribute almost two-thirds of the cancer risk-weighted emissions. Acrolein and formaldehyde appear significant when considering risk-weighted emissions for chronic and acute effects. Phosphorous also appears significant to risk-weighted emissions for chronic effects. On-road mobile sources and aircraft are the two most important source categories for non-cancer risks, comprising 40% of all chronic risk-weighted emissions and almost 80% of acute risk-weighted emissions.

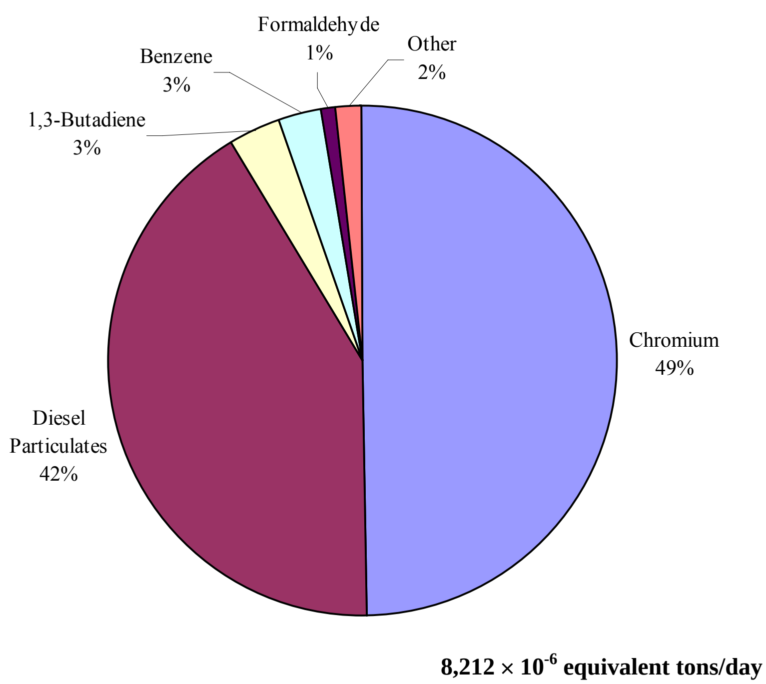


Figure 4-1. Cancer risk-weighted emissions by pollutant.

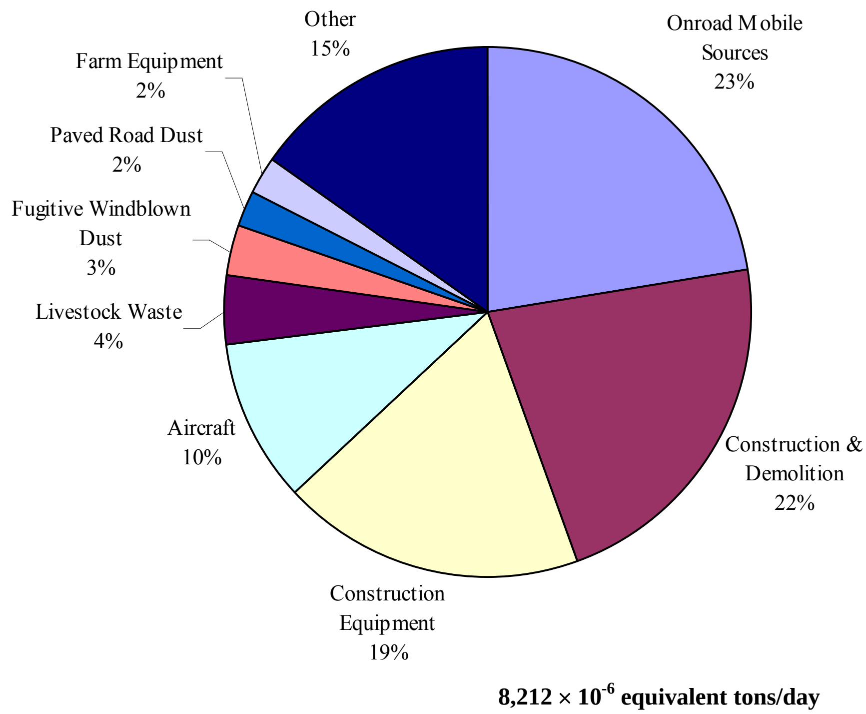


Figure 4-2. Cancer risk-weighted emissions by source category.

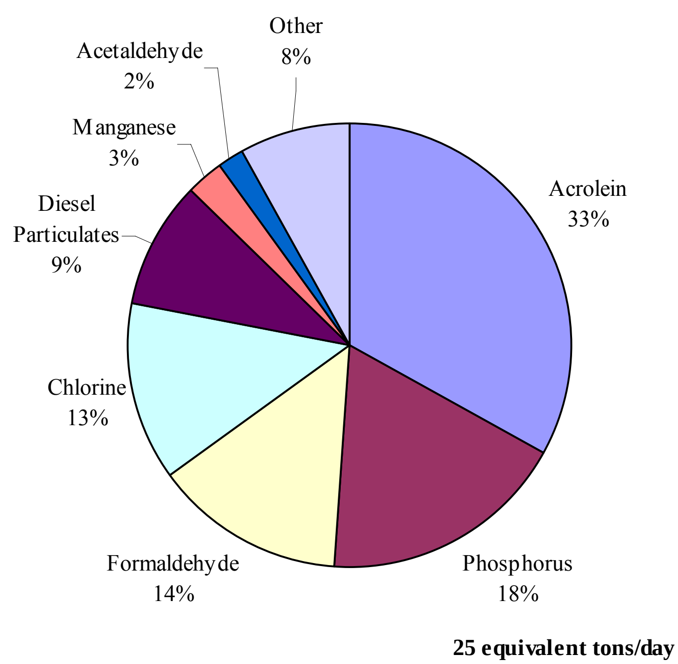


Figure 4-3. Chronic risk-weighted emissions by pollutant.

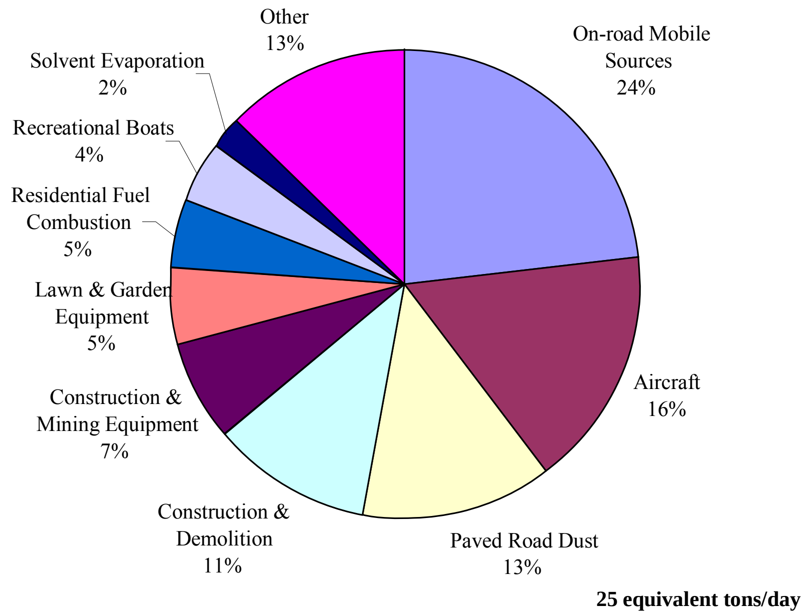


Figure 4-4. Chronic risk-weighted emissions by source type.

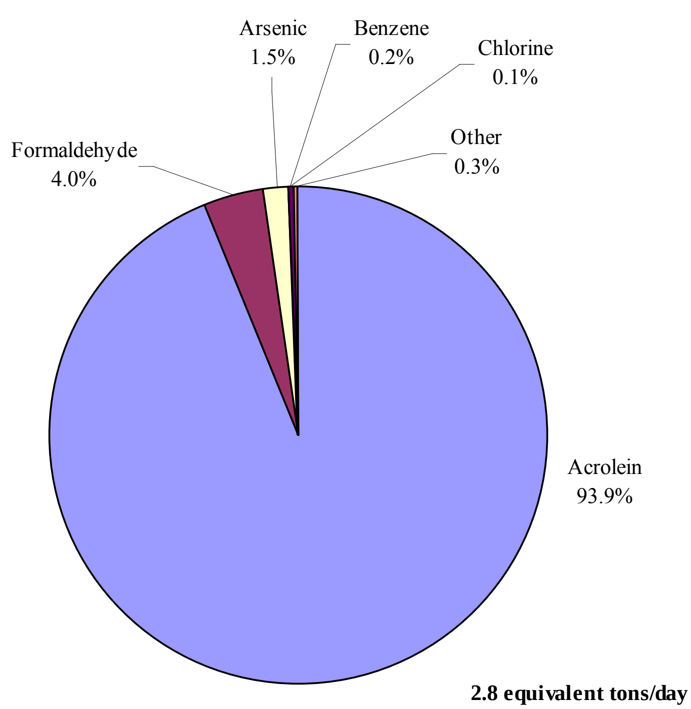


Figure 4-5. Acute risk-weighted emissions by pollutant.

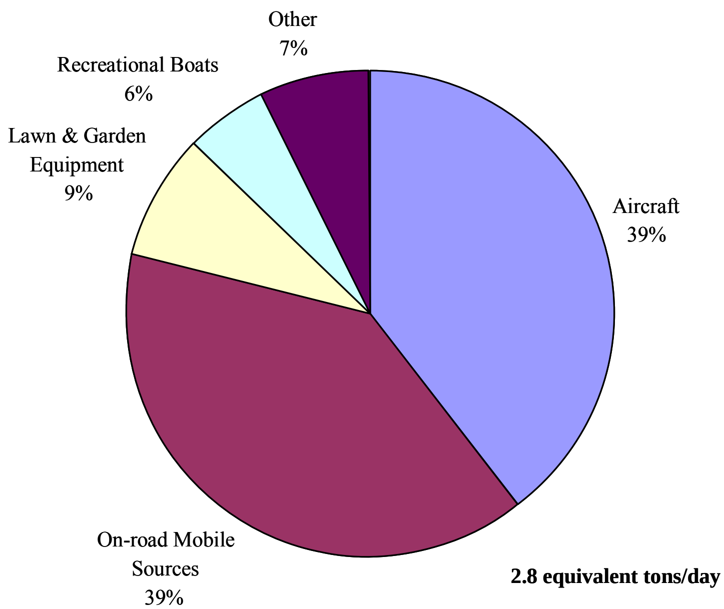


Figure 4-6. Acute risk-weighted emissions by source category.

Geographic distributions of emissions for several of the TACs prominent in Figures 4-1 through 4-6 are illustrated in **Figures 4-7 through 4-14**, including chromium, diesel particulate matter, acrolein, phosphorus, chlorine, MTBE, formaldehyde, and toluene. For visual comparison, **Figure 4-15** illustrates population density in the Bay Area (Environmental Systems Research Institute, 2002).

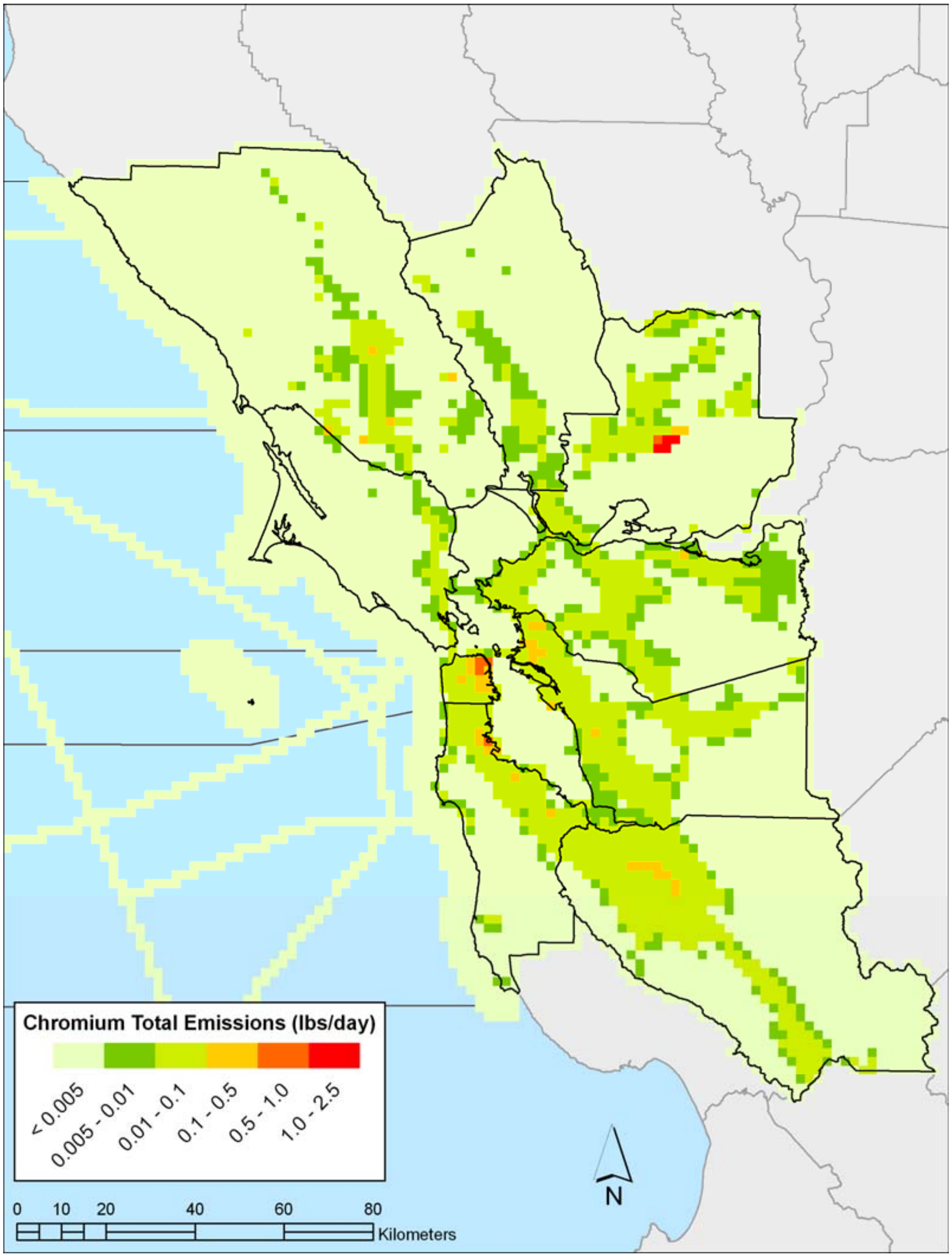


Figure 4-7. Emission density plot of chromium emissions.

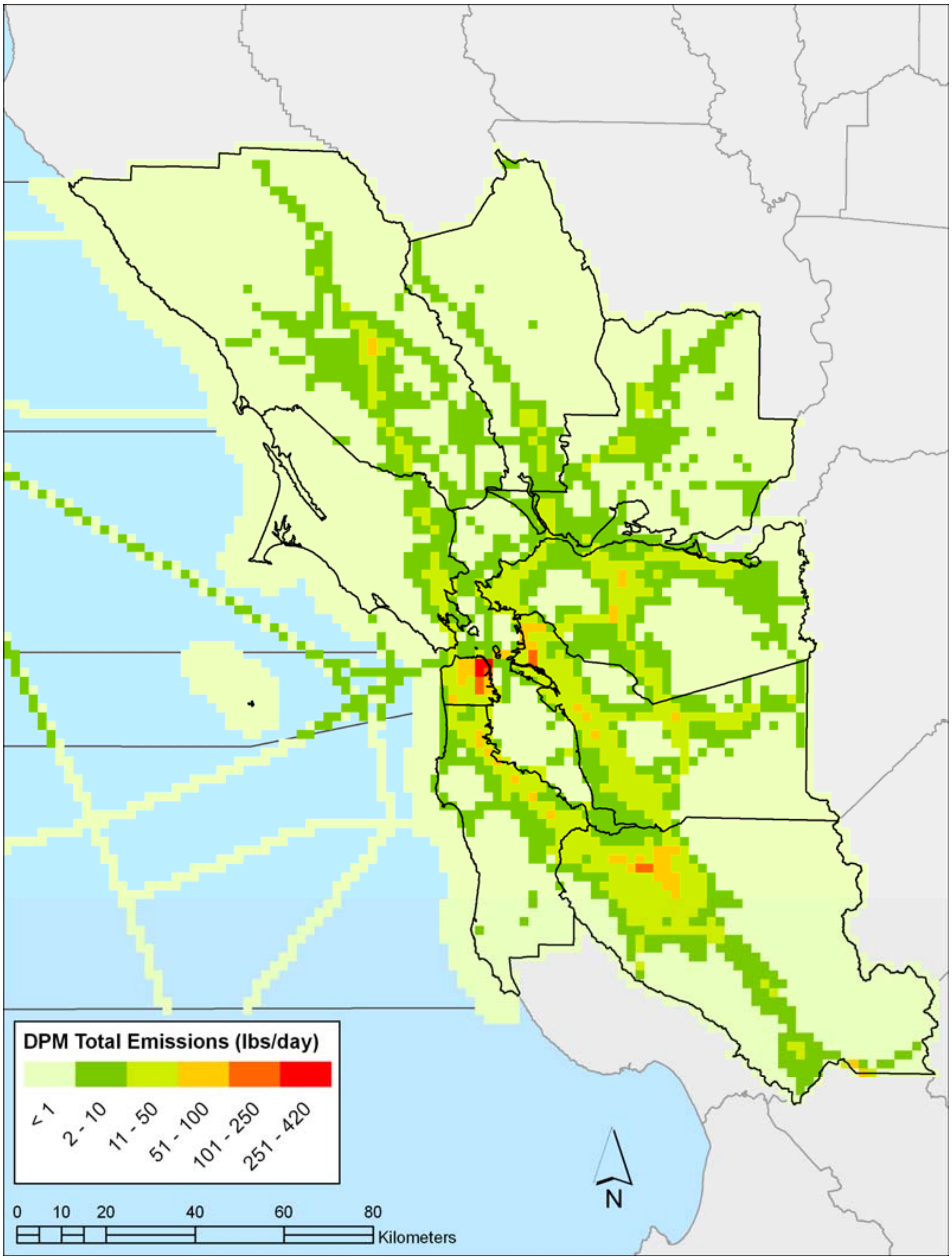


Figure 4-8. Emission density plot of diesel particulate emissions.

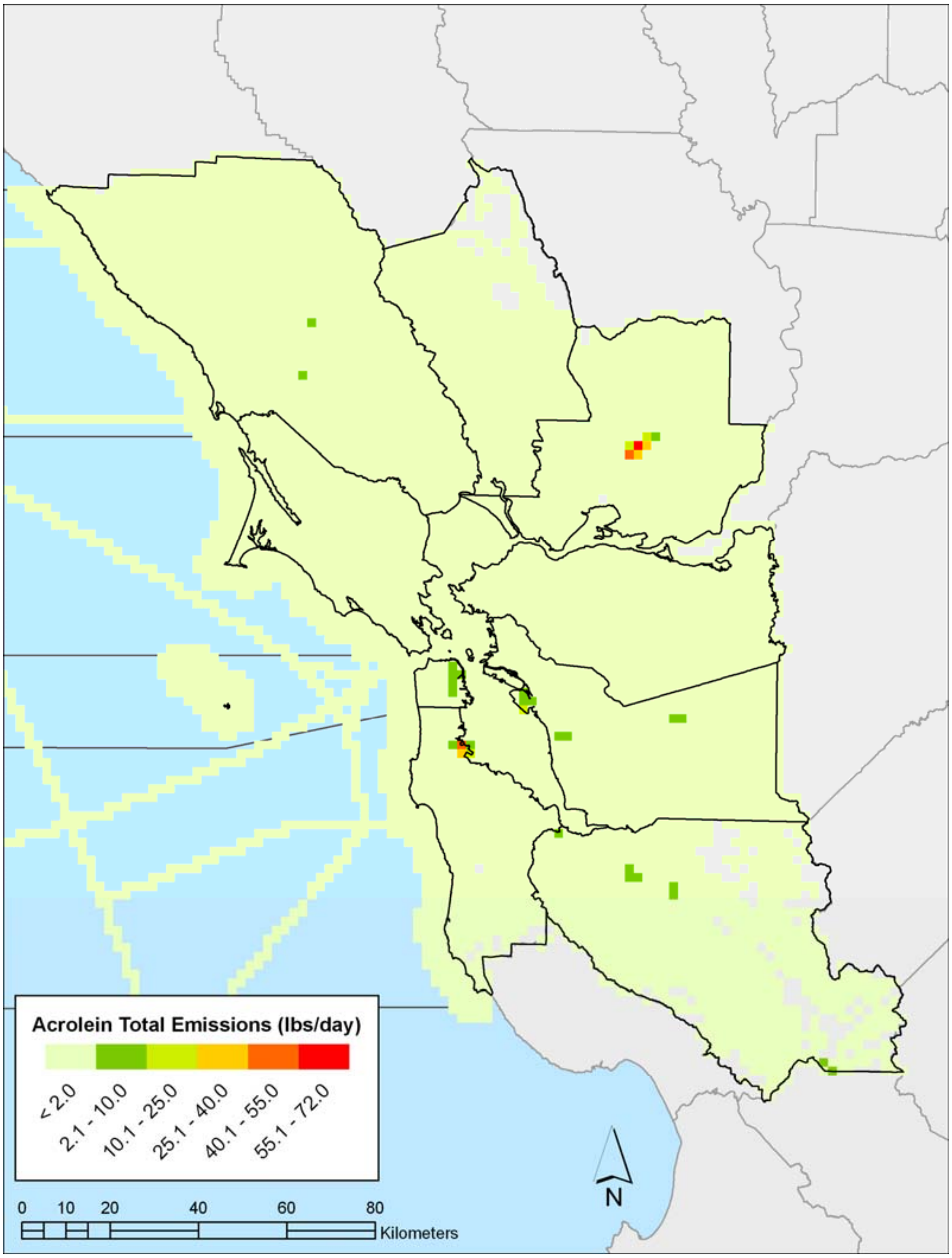


Figure 4-9. Emission density plot of acrolein emissions.

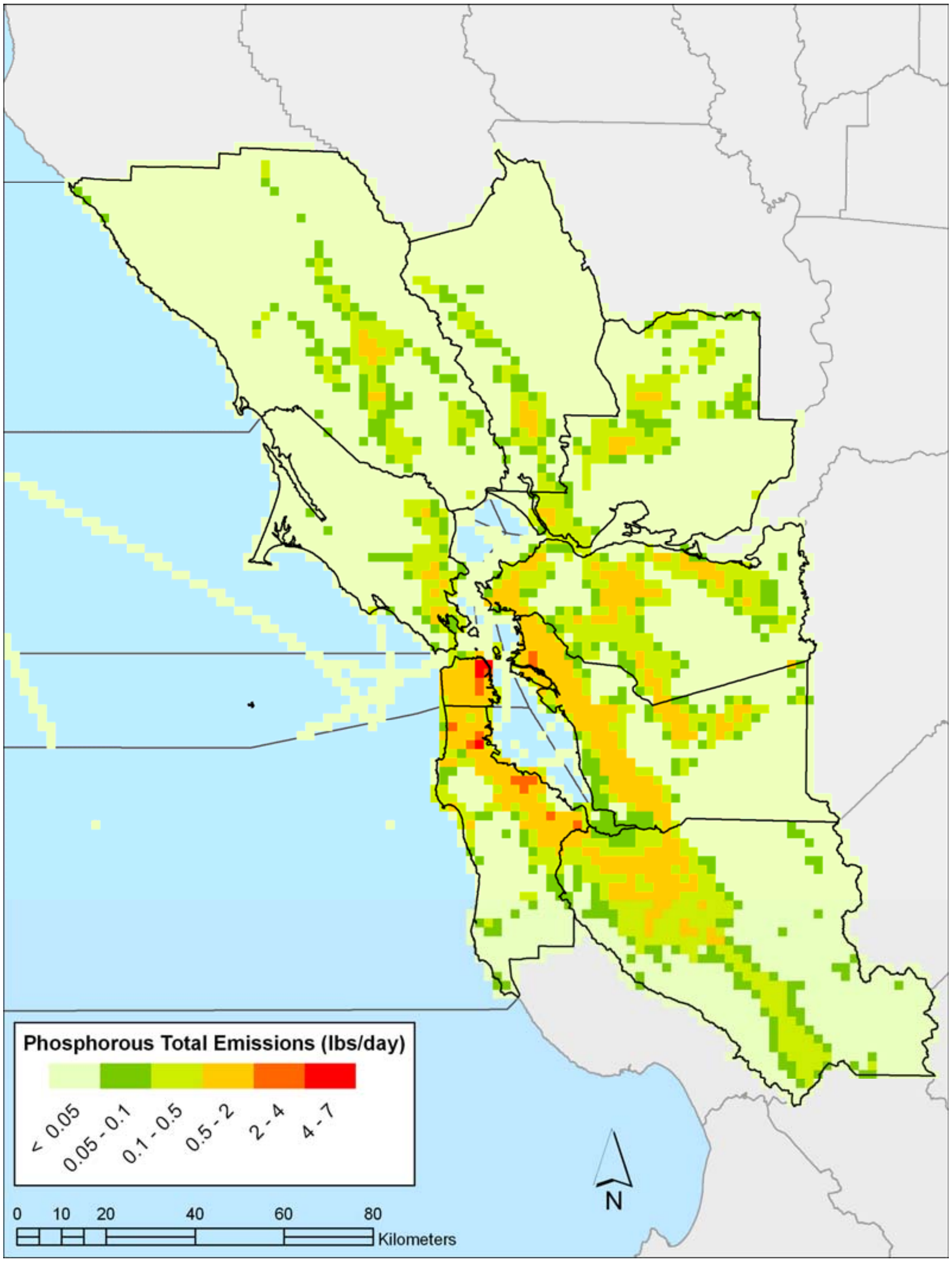


Figure 4-10. Emission density plot of phosphorous emissions.

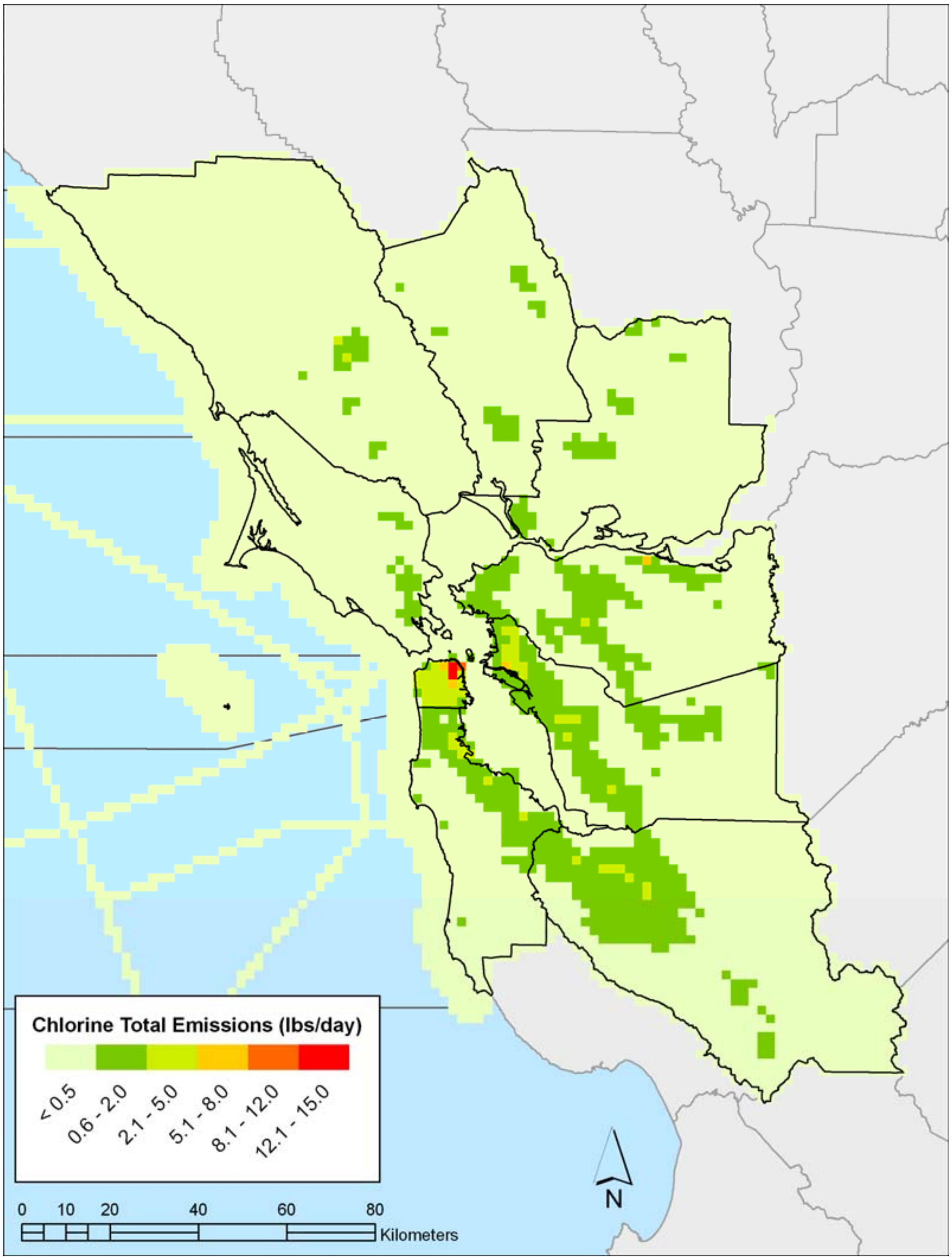


Figure 4-11. Emission density plot of chlorine emissions.

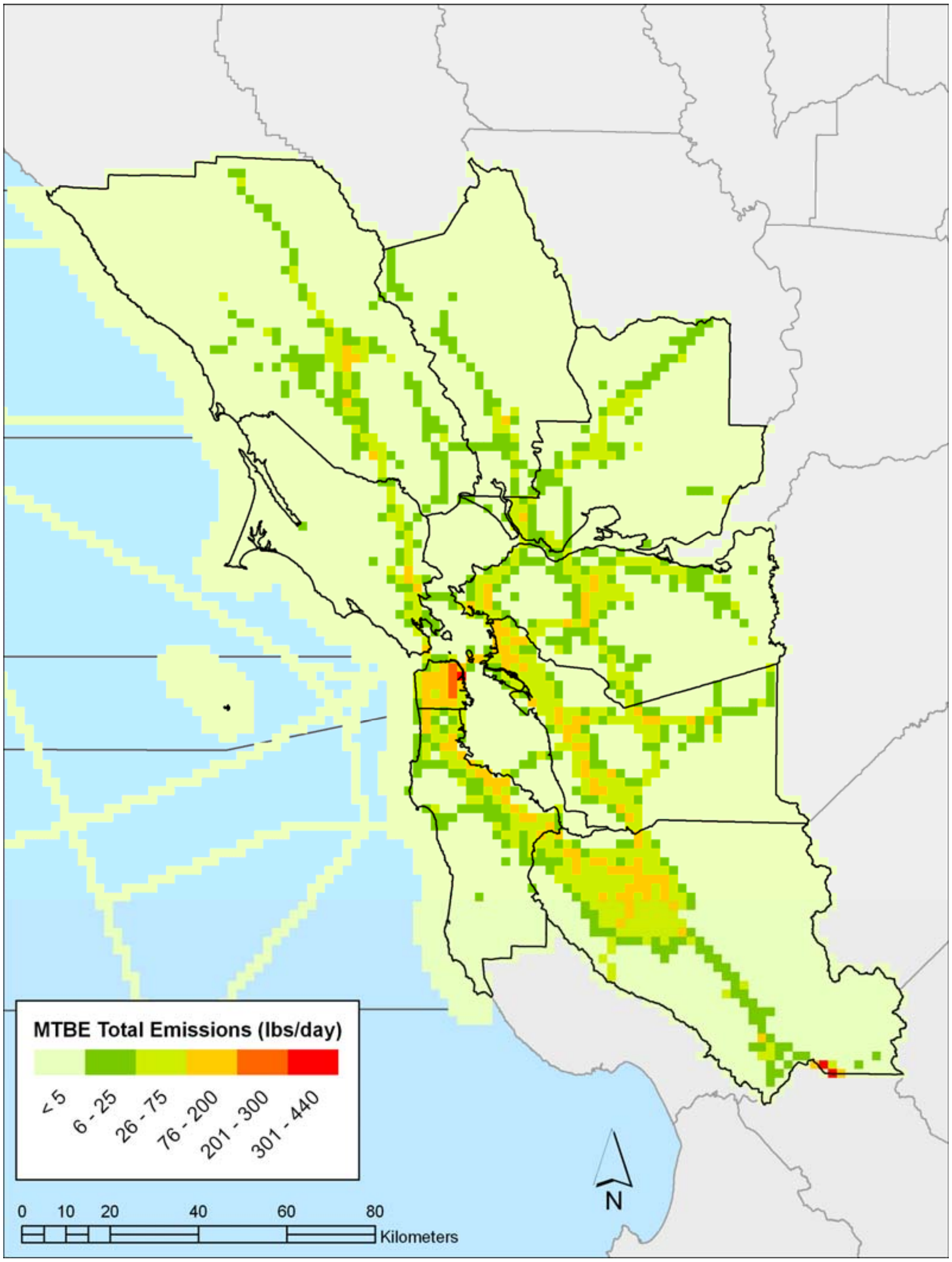


Figure 4-12. Emission density plot of MTBE emissions.

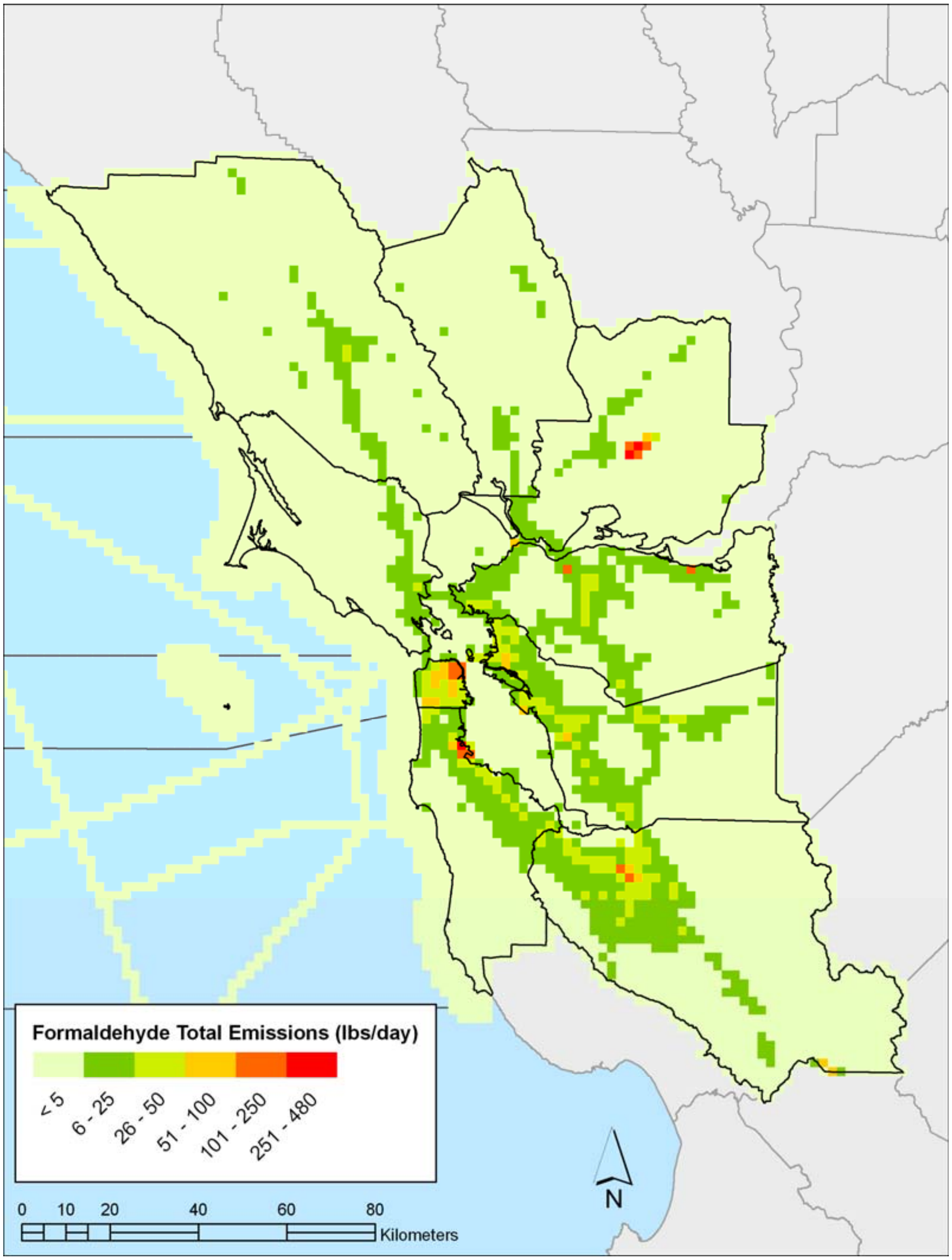


Figure 4-13. Emission density plot of formaldehyde emissions.

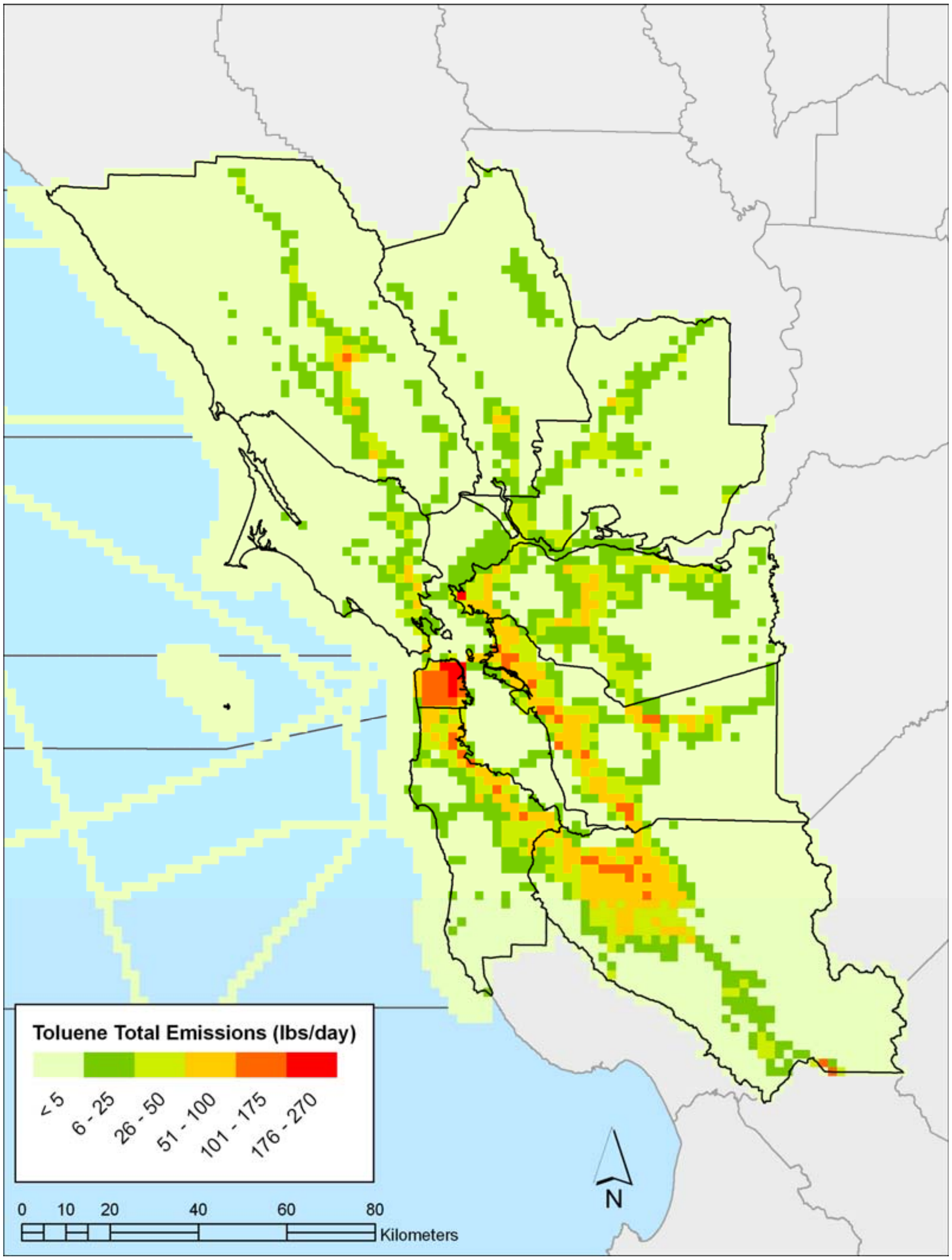


Figure 4-14. Emission density plot of toluene emissions.

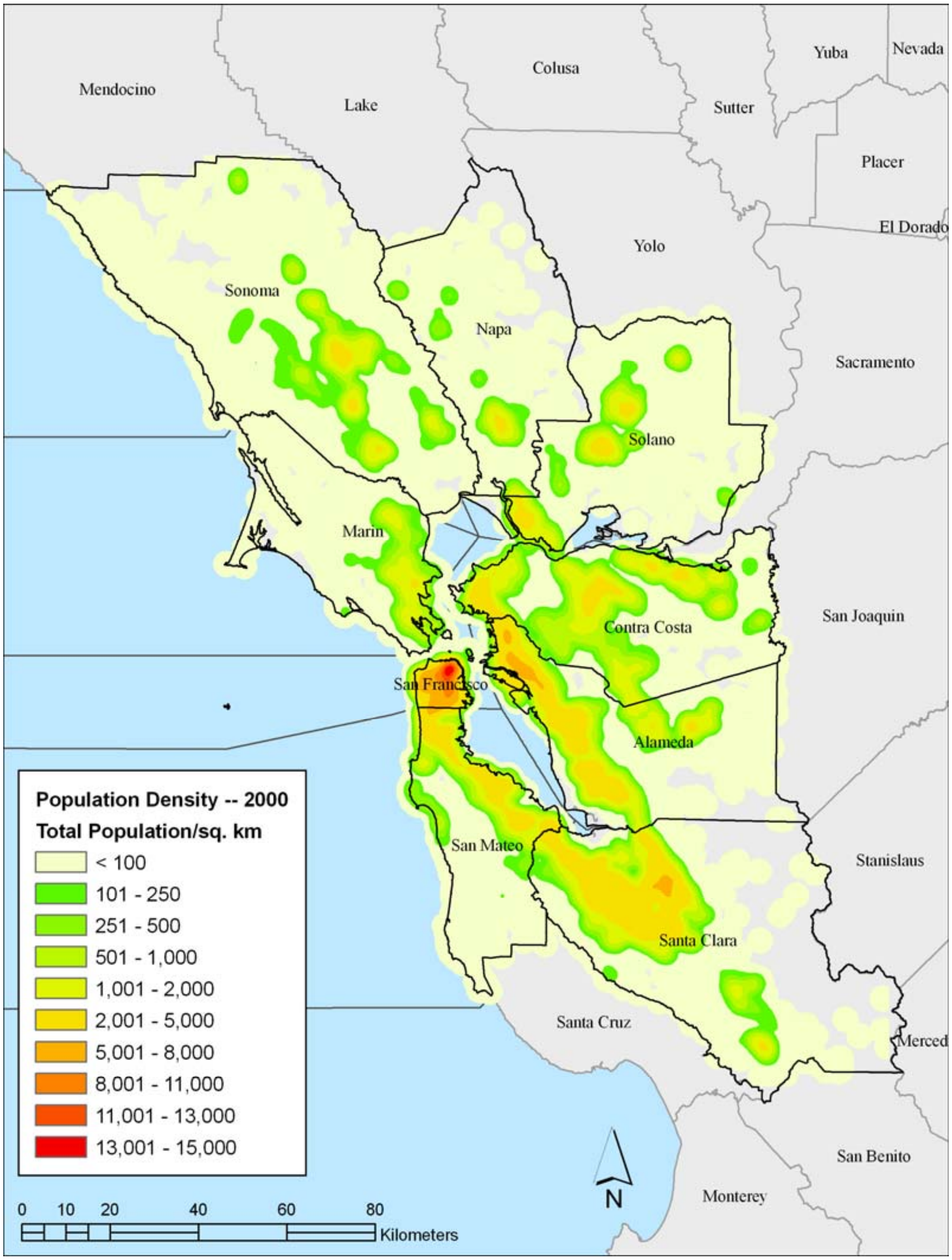


Figure 4-15. Population density in the San Francisco Bay Area.

5. CONCLUSIONS AND RECOMMENDATIONS

The TAC inventories summarized in this document represent the first step of inventory development for the CARE program. These screening-level inventories are useful in prioritizing further inventory development activities, conducting preliminary exposure modeling runs, and selecting a study community in the Bay Area. This section characterizes general aspects of the TAC inventories and provides recommendations for continued inventory development.

The TAC inventories prepared to date should be treated as "working versions under development". They were prepared from readily available, pre-existing information sources—an approach that was both time- and cost-efficient. A more technically rigorous approach to inventory development—though requiring more effort and resources—is a "bottom-up" approach. Bottom-up inventory development involves gathering observations of emissions or supporting data directly from emissions sources. Examples of bottom-up methods include direct measurements of emissions at their sources; surveys of emissions-related activities by telephone, mail, or in person; or measurement of emissions-producing activities with monitoring devices, such as automated daily traffic counters or meters that measure fuel consumption.

The TAC inventories harbor some strengths, but also several potentially important weaknesses. Strengths of the inventories include

- TAC emissions for on-road mobile sources, which were estimated using EPA-recommended methods but were improved with California-specific chemical speciation data;
- TAC emissions for point sources, which were directly reported and were likely based on bottom-up approaches; and
- Spatial allocations of emissions, which produced inventories that are accurate and spatially well-resolved relative to the resolution of the modeling grid (2 km × 2 km).

However, some TACs are likely omitted from the inventory because they are not components of TOG or PM_{10} (e.g., hydrofluoric acid) or because they are infrequently measured for chemical speciation profiles. Two examples of omitted TACs—quinoline (most often emitted by combustion sources) and hydrazine (most often released by certain industrial manufacturing processes)—are among the EPA's designated urban air toxics, a listing of 33 priority TACs (plus diesel particulate matter [DPM] and coke oven emissions), which are considered to pose significant health risks in urban areas of the U.S. Others—which may or may not be important in the San Francisco Bay Area—include radionuclides (usually from natural sources and controlled medical or testing uses); titanium tetrachloride (from titanium metals manufacturing); and hydrochloric acid, hydrofluoric acid, and sulfuric acid (most often emitted by certain manufacturing processes and refining).

In addition, the chemical speciation profiles that were applied to estimate TAC emissions from pre-existing emission inventories of TOG and PM_{10} contain significant uncertainties. Due to the limitations of available chemical analysis techniques, some of the speciation profiles include large reported proportions of unknown or unidentified species—occasionally as much as 50% to 80% of the total mass—contributing to total TOG or PM emissions. Measurements of source-specific chemical speciation profiles are expensive; therefore, the number of measurements for each source category and TAC are fairly limited in number and may contain inaccuracies or large errors. We are particularly concerned about the large proportion of chromium attributed to the ARB's source profile for construction and demolition activities (as noted in Section 5.1). However, in general, these issues with speciation profiles are difficult and expensive to resolve and will likely take many years of gradual efforts to address.

Most importantly, a large degree of uncertainty was introduced by conservatively assuming that chromium compounds (unspecified), which are associated mostly with fugitive dust emissions, are emitted in their most highly toxic form: chromium (VI) rather than chromium (III). As a result of this conservative assumption, cancer risk-weighted emissions may be overestimated by 4,100 x 10^{-6} equivalent tons per day (or 50% of total cancer risk-weighted emissions). Chronic risk-weighted emissions may be overestimated by 3.0 equivalent tons per day (or 12% of chronic risk-weighted emissions). Therefore we recommend conducting TAC assessments with two different scenarios for comparison—including and excluding emissions of unspecified chromium compounds (CAS no. 7440473)—in order to assess the effects of this large uncertainty. In addition, we recommend further research into the relative proportions of chromium (III) and chromium (IV) emitted in fugitive dust.

5.1 RECOMMENDATIONS FOR CONTINUING DEVELOPMENT OF THE TOXIC AIR CONTAMINANT EMISSION INVENTORIES

The following strategies are suggested to begin addressing some of the weaknesses in the TAC inventories:

- To the extent feasible, use bottom-up methods for emission inventory development beginning with the highest priority TACs and source types. Prioritize TACs that seem most likely to pose health risks in the San Francisco Bay area for further emission inventory development. We suggest prioritizing TACs that are listed among the EPA's designated urban air toxics or that appear prominently in Figures 4-1, 4-3, and 4-5. Identify source categories likely to emit prioritized TACs, such as those shown in Figures 4-2, 4-4, and 4-6 or identified in EPA guidance documents (U.S. Environmental Protection Agency, 2005; 1998, Appendix I).
- Use readily available (or, if possible, bottom-up methods) to estimate emissions of TACs that are currently omitted from the inventories: quinoline, hydrazine, radionuclides, titanium tetrachloride, hydrochloric acid, hydrofluoric acid, and sulfuric acid.
- Correct the coordinates of point sources that were located according to zip code centroids—137 facilities of 3,359 in all—beginning with those estimated to emit the largest quantities of prioritized TACs. A list of point sources located by zip code centroids is provided in Appendix B.
- The ARB maintains a database of TAC emissions reported for its AB 2588 Air Toxics Hot Spots Program (California Air Resources Board, 2001b). In a few cases, 1996 emissions reported in the AB 2588 database for point sources in the Bay Area greatly exceeded year-2000 emissions included in the BAAQMD's TAC inventory for point sources. These differences may accurately represent reduced levels of TAC emissions;

however, it may be worthwhile to verify that this is the case. The AB 2588 database reported 972 tons per year (tpy) of methylene chloride emitted from point sources in the Bay Area (compared to 59 tpy included in the BAAQMD inventory); 5 tpy of manganese (compared to 0.6 tpy); and 110 pounds per year of beryllium (compared to 1 pound per year).

- We recommend investigating the ARB's recommended PM_{10} speciation profile for construction and demolition because it appears to contain a relatively large fraction of chromium—roughly 10 times larger than the chromium contents of paved and unpaved road dust. However, we acknowledge that chromium VI is a recognized compound of concern in Portland cement (Klemm, 1994); therefore, the high chromium content of construction and demolition dust may be justified. If the chromium content of PM_{10} emissions from construction and demolition activities were reduced by a factor of 10, then total cancer risk-weighted emissions in the BAAQMD's TAC inventory would decline 20% to 6,586 x 10^{-6} equivalent tons per day. Of the reduced cancer risk-weighted emissions, 3% would be attributable to fugitive dust from construction and demolition dust and 33% would be attributable to other sources of chromium.

5.2 RECOMMENDATIONS FOR REVIEWING THE CRITERIA POLLUTANT INVENTORIES

The BAAQMD's TOG and PM_{10} emission inventories, which were the basis of the TAC inventories, previously underwent a thorough quality assurance and quality control (QA/QC) review to support of the goals and objectives of the Central California Ozone Study (CCOS).[2] Therefore, the scope of the BAAQMD's TAC inventory development project included only cursory QA/QC of the inventory files as received from BAAQMD. No significant problems were found and the inventories were used "as-is". However, in the course of reviewing the charts and figures presented in this report, we noticed a few unusual features, which—though very minor—might warrant investigation.

- A small "hot spot" (about 5 grid cells) of emissions from on-road mobile sources was noted in southeast Santa Clara County. This hot spot is located along a rural stretch of Highway 152 east of Gilroy and north of Hollister. The emissions appear unusually large for a rural area. Approximately 2% of total estimated on-road mobile emissions in the BAAQMD counties are within this hot spot.
- Emissions from on-road mobile sources along Highway 1 in San Mateo County appear to be slightly mis-located approximately 2 km too far west (over Pacific waters).
- We considered whether emissions estimated for ships and commercial boats might be too low for an area with active ports. We verified the reasonableness of the emissions by comparing the BAAQMD's emission inventories to three other inventories—those for Baton Rouge, Louisiana; Houston-Galveston, Texas; and Los Angeles, California. These three port cities were selected for the comparison because we are familiar with the underpinnings of their inventories and are highly confident about the validity of the estimated emissions for commercial marine vessels. In these three cities, commercial

[2] More information about the CCOS is available at http://www.arb.ca.gov/airways/ccaqs.htm.

marine vessels were estimated to emit 10% to 24% of total $PM_{2.5}$ from all non-road mobile sources: 24% in Baton Rouge, Louisiana; 10% in Houston-Galveston, Texas; and 16% in the SCAQMD (Reid et al., 2004; California Air Resources Board, 2001a). The BAAQMD's inventory attributes 11% of total non-road $PM_{2.5}$ emissions to commercial marine vessels, which is within the range of values estimated for the other cities.

6. REFERENCES

California Air Resources Board (2001a) The 2001 California almanac of emissions and air quality. Planning and Technical Support Division: Sacramento, California, 2001.

California Air Resources Board (2001b) California toxics inventory. Available on the Internet at <http://www.arb.ca.gov/toxics/cti/cti.htm> last accessed July 15, 2005.

Environmental Systems Research Institute I. (2002) Census 2002 TIGER/line data. Available on the Internet at <http://www.esri.com/data/download/census2000_tigerline/index.html> last accessed July 13, 2005.

Funk T.H., Stiefer P.S., and Chinkin L.R. (2001) Development of gridded spatial allocation factors for the State of California. Technical memorandum prepared for the California Air Resources Board, Sacramento, CA, by Sonoma Technology, Inc., Petaluma, CA, STI-900201/999542-2092-TM, July.

Klemm W.A. (1994) Hexavalent chromium in Portland cement. *Cement, Concrete, and Aggregates* **16** (1), 43-47.

Reid S.B., Sullivan D.C., Penfold B.M., Funk T.H., Tamura T.M., Stiefer P.S., Raffuse S.M., and Arkinson H.L. (2004) Emission inventory development for mobile sources and agricultural dust sources for the Central States. Final report prepared for The Central States Air Resource Agencies and The Central Regional Air Planning Association, Oklahoma City, OK, by Sonoma Technology, Inc., Petaluma, CA, STI-903574-2611-FR, October.

U.S. Environmental Protection Agency (1998) Handbook for air toxics emission inventory development. Office of Air Quality Planning and Standards, EPA-454/B-98-002, November.

U.S. Environmental Protection Agency (2005) Locating and estimating (L&E) documents. Available on the Internet at <http://www.epa.gov/ttn/chief/le/> last accessed July 15, 2005.

B.3 SUITABILITY MODELING

Following is a draft manuscript about predicting areas of high diesel particulate matter emissions in Phoenix, Arizona. The manuscript details the methods used in suitability modeling. In this case, the goal was to find ideal locations for monitor placement for measuring maximum emissions of diesel particulate matter. These methods apply to network assessment for other monitoring objectives.

Predicting Areas of High Diesel Particulate Matter Emissions in Phoenix, Arizona Using Spatial Analysis Techniques

Bryan M. Penfold, Steven G. Brown, Hilary R. Hafner, and Tami H. Funk
Sonoma Technology, Inc.
1360 Redwood Way, Suite C
Petaluma, California

ABSTRACT

In this study, geographic information system (GIS)-based suitability modeling was performed to identify areas throughout Phoenix, Arizona, where DPM emissions are likely to be high and emanate from a mix of source types. Maps of emissions sources, emissions activity data, and meteorology were combined within a GIS model to produce a composite map identifying regions where DPM emissions are likely to be high. The results of the GIS model were compared to (a) the locations of existing ambient toxics monitoring sites in Phoenix and (b) the spatial distribution of population in the region. The results indicate that two of the existing monitoring sites are located in areas where DPM emissions are predicted to be high; however, when meteorology is considered, one of the sites is located upwind of a predicted high DPM area. When population density is considered, two monitoring sites in central Phoenix are located in areas of high total population density. Overall, this analysis illustrated the usefulness of suitability modeling as an aid in selecting monitoring locations to meet specific objectives.

INTRODUCTION

Diesel Particulate Matter (DPM) is part of a complex mixture that makes up diesel exhaust. Diesel exhaust is commonly found throughout the environment and is estimated by the U.S. Environmental Protection Agency's (EPA's) National Scale Assessment (NSA) to contribute to human health risk[1] and can cause acute and chronic health effects.[2-4] It is also a significant contributor to $PM_{2.5}$ concentrations and regional haze.[5-8] As such, DPM has been the focus of ambient monitoring and long-term epidemiological studies. Diesel exhaust is emitted by a broad range of diesel engines including on-road diesel trucks, locomotives, marine vessels, and heavy-duty equipment. These various sources emit different amounts of DPM, and are often spatially dispersed within an urban area. DPM concentrations are highest and have the best correlation with respiratory distress near the areas of highest diesel usage.[9-11]

DPM cannot be directly measured; elemental carbon (EC) or black carbon (BC) measurements are often used as a surrogate although measurements of EC or BC alone are insufficient to quantify diesel contributions. Different sources of DPM emit different amounts of EC or BC relative to organic carbon, and analytical methods for EC and BC differ,[12,13] making integrated spatial monitoring of DPM difficult. Therefore, novel approaches must be developed to determine areas of DPM influence, to assist in identifying suitable monitoring locations to target DPM, and to provide data which may be useful to assess whether sensitive populations may be adversely impacted by DPM.

Suitability modeling analysis has been used for several studies predicting atmospheric benzene and EC hot spots within an urban area,[14,15] deposition of various pollutants in marine environments[16-18] as well as to identify possible locations for placing monitors.[15]

The objective of this work was to use geographic information system (GIS) technology to identify areas within the Phoenix region where DPM emissions are likely to be high and to identify locations that are potentially suitable for placing toxics monitors to better measure DPM. Figure 1 illustrates the study domain and locations of existing long-term toxics monitoring sites in Phoenix.

METHODS

Overview

Suitability modeling is a method for identifying suitable monitoring locations based on specific criteria. For example, suitability modeling can be used to determine possible locations for new air quality monitoring sites based on criteria such as emissions source influence, proximity to populated places, urban or rural land use, site accessibility, etc. The idea is that map layers representing these important criteria can be compiled and merged to develop a composite map representing the combination of important criteria for a defined area. Furthermore, each of the map layer inputs can be assigned a weighting factor based on the relative importance of each layer in the overall suitability model. For example, when determining suitable locations for placing a new air quality monitor, each of the important criteria can be prioritized in terms of its relative importance. If the monitoring objective is to measure air quality in densely populated places then a map layer representing population density would be given priority, and a corresponding high weighting factor, in the overall model and the resulting suitability map output would favor areas of high population density.

The Environmental Systems Research Institute (ESRI) ArcGIS software, Spatial Analyst, was used for this analysis. Spatial Analyst is a raster- or grid-based software package that provides a platform for developing and manipulating gridded data. Spatial Analyst can be used to develop suitability models that produce maps highlighting “suitable” geographic regions based on defined model criteria and weighting schemes.

The following general steps are used to develop a suitability model as illustrated by Figure 2:

1. *Define the objective.* The first step in developing a suitability model is to determine the objective or scenario to be modeled such as identifying where DPM emissions are likely to be high in the Phoenix area. Suitability maps help identify regions that are suitable for a specific purpose. For example, a suitability map could be used to determine where to place new schools or hospitals based on many criteria including population, access, and service needs.

2. *Identify and acquire data sets*. The second step in suitability modeling is to determine which data sets are required for the analyses. This is achieved by identifying geographic features and data that are important to include in the model.

3. *Derive data sets*. The third step in building a suitability model is to derive new, gridded data from the input data sets. These new data may consist of distance contours or density plots based on the original map feature data. The data are then reclassified to create a common data scale. For example, if the distance from a particular geographic feature is a required input to the suitability model, distance contours can be derived from the locations of geographic features.

4. *Weight and combine data sets*. The final steps in suitability modeling are to determine the relative importance of each data set, to weight them accordingly, and to combine the data sets to produce a suitability map. Weighting the data sets defines the extent to which each data set will influence the model results.

The following three steps were performed to identify areas in the Phoenix region likely to be influenced by DPM:

1. Assess the emission inventory to determine the predominate sources of DPM in the region and the best available data to represent the spatial pattern of the identified emissions sources in the Phoenix region.

2. Acquire and process the spatial data (map layers) required for the analysis.

3. Develop the suitability model to predict areas that are likely to have high DPM emissions.

A key feature of this GIS-based approach is that the analysis can be performed without actual monitored data. Results can be improved by obtaining finer detail for the spatial layers, adding more layers of pertinent information, and improving the emissions and activity information regarding DPM.

Step 1: Emission Inventory Assessment

The first step in this analysis was to assess the emission inventory for the Phoenix area and to identify the important sources of DPM in the region. The 1999 U.S. EPA National Emission Inventory (NEI) was acquired and processed for Maricopa and Pinal counties. Diesel sources tend to emit substantial levels of $PM_{2.5}$; therefore, the NEI for Maricopa and Pinal counties was assessed to determine the predominant sources of $PM_{2.5}$. Table 1 lists the major sources, emissions, and percent contribution to total $PM_{2.5}$ emissions for Maricopa and Pinal counties.

Figure 3 details the source breakdown for $PM_{2.5}$ emissions in Maricopa and Pinal counties. Approximately 12% (4,372 tons/year) of total $PM_{2.5}$ emissions in these counties are from on- and non-road mobile sources. Area and point sources are responsible for about 86% (32,851 tons/year) and 2% (694 tons/year) of total $PM_{2.5}$ emissions, respectively. As indicated in Table 1 and Figure 3, the top two sources of $PM_{2.5}$ are road construction activities and open burning. Road construction activities include both exhaust and dust emissions. Open burning and wildfire emissions were not considered important for this analysis because they are not a significant source of DPM. Moreover, open burning and wildfires tend to occur in the rural areas outside of Phoenix and the focus of this work is on more densely populated urban areas.

Because diesel emissions are the primary focus of this analysis, sources of dust such as road construction and fugitive road dust were also excluded when possible. In some cases, sources of dust and exhaust were combined into one source category; emissions for these combined categories were included in the analysis.

To help quantify the sources of $PM_{2.5}$ listed in Table 1 in terms of their potential contribution to DPM, PM speciation profiles were acquired from the California Air Resources Board (ARB) and EPA's SPECIATE Database[19,20] to determine the relative mass fractions of EC from each of the major sources as a surrogate for diesel emissions.[6,21,22] The speciation profiles were multiplied by the mass of emissions for select sources to arrive at an approximate mass for EC by source type. Table 2 lists the EC contributions for a subset of sources from Table 1. As shown in Table 2, on-road heavy-duty diesel vehicles have the highest relative levels of EC, followed by diesel construction, and on-road gasoline vehicles.

Step 2: Data Acquisition and Processing

Because we are interested in identifying locations where DPM emissions are likely to be high, we must be able to spatially characterize the distribution of emissions for each major DPM source category. For example, a map of roadways and associated traffic volumes for heavy-and light-duty vehicles could be used to characterize the spatial distribution of emissions from on-road mobile sources. A less straightforward example is construction equipment. Because construction equipment is a mobile source and the exact locations of emissions releases are less known, surrogate map layers can be used to represent emissions from these source types. For example, maps indicating areas of new development and construction could be used as a surrogate for diesel construction equipment emissions.

An important aspect of this analysis is assessing the proximity of identified areas where DPM is likely high to population density. Geophysical land features and meteorology are also important to include in the model because they influence the dispersion of emissions.

Several sources of spatial data were identified and assessed for use in this analysis. Table 3 lists each of the major emissions sources and the corresponding map layer that was assigned to represent the spatial distribution of emissions in the suitability model. Population data by Census block for 2000 were acquired from the U.S. Census Bureau. Digital elevation model (DEM) topography data were acquired from the United States Geological Survey (USGS).

Spatial data representing on- and non-road mobile sources were acquired from a variety of sources. Road network maps containing annual average daily traffic (AADT) data for heavy- and light-duty vehicles in the Phoenix area were obtained from the Arizona Department of Transportation (ADOT). AADT is an indicator of the relative on-road mobile source activity, and corresponding emissions levels, in the Phoenix area. Figure 4 shows the road network map and the AADT data for (a) light-duty and (b) heavy-duty vehicles by road segment. The locations of airport, rail lines, and transportation distribution facilities, also sources of DPM, were obtained from the U.S. Bureau of Transportation Statistics (BTS). We did not have airport, rail or transportation distribution facility activity information; this type of emissions data could be incorporated in the model if available.

Spatial surrogate data were acquired to represent the spatial pattern of DPM emissions from non-road mobile source equipment including construction equipment and lawn and garden equipment. Because the majority of construction equipment usage occurs during the development and construction of residential and commercial buildings, maps of residential and commercial development completions from 2000-2003 were obtained from the Maricopa Association of Governments (MAG). The square footage of residential development and the acreage of commercial development were used as indicators of the relative emissions activity for construction equipment. To spatially represent the distribution of emissions from lawn and garden equipment, land use data identifying large irrigated grass areas (i.e., golf courses and cemeteries) were obtained from MAG. Figure 5 identifies the locations of golf courses, cemeteries, and large development areas in the Phoenix area.

Emissions data for large point sources were obtained from the EPA's Air Quality System (AQS), and were used to map the magnitude of $PM_{2.5}$ emissions in the Phoenix area. Figure 6 identifies the location and emissions contributions of large industrial facilities that emit $PM_{2.5}$ in the Phoenix area. Point source locations were used to investigate the impact other $PM_{2.5}$ sources would have on effectively assessing areas of potentially high DPM concentrations. Digital elevation model (DEM) data were acquired to produce a three-dimensional visualization of the regional topography. DEM data were also used to characterize the potential topographical influence on meteorology and the distribution and transport of emissions.

One unique feature of this analysis was the attempt to account for meteorological influences in the suitability analysis. For example, not only do we expect DPM concentrations to be higher closer to an emission source, we also expect concentrations to be higher in areas downwind of the source. Wind speed and direction data were acquired for 12 monitoring sites within the Phoenix area from the Central California Air Quality Studies (CCAQS) database. Using the CCAQS data, annual average gridded wind fields were developed using data from meteorological stations to represent the predominant wind direction throughout the region. Figure 7 is a wind rose plot of the annual average wind speed and direction for Phoenix based on data from 2001-2003. As shown in Figure 7, the colored coded bars represent the percent and speed of wind from directions ranging 1 to 360 degrees.

Population data were acquired from the 2000 U.S. Census and were used to create maps of the regional population distribution. The population data were used to investigate the placement of existing monitors relative to total population (Figure 8). As shown in Figure 8, central Phoenix has the highest total population density in the region. Other population groups could be investigated, such as different age or income ranges, depending on the monitoring or analysis objectives.

Step 3: Suitability Model Development

Two model scenarios were defined to examine the spatial distribution of DPM emissions: (1) development of a composite map to represent the spatial distribution and density of DPM emissions based on the locations of DPM sources (hot spots) and (2) proximity of total population to DPM sources. Each of the model scenarios were developed both including and excluding meteorological effects (wind speed and direction). The model scenarios were developed by assessing each emissions source and its relative contribution of EC emissions (used in this analysis as a surrogate for DPM emissions). Each map layer representing the spatial pattern of emissions was assigned a weighting factor to determine its contribution to the outcome of the overall suitability model depending on the objective of each model scenario.

The first step in developing the modeling scenarios was to determine which source types contribute significantly to EC emissions. As shown in Figure 3, area sources (which include non-road construction equipment) are the largest contributor to total $PM_{2.5}$ emissions. On-road mobile sources are the next largest contributor and point sources only contribute 2% to total $PM_{2.5}$ emissions. As noted in Table 2, the highest EC contribution comes from heavy-duty highway diesel vehicles, followed by diesel construction, and gasoline vehicles.

The second step in developing the modeling scenarios was to develop a weighting factor for each map layer based on the EC contributions corresponding to the emissions source represented by the map layer. Table 4 summarizes the relative EC contributions corresponding to each map layer and the assigned weighting factor. For example, EC contributions from diesel construction and mining equipment, diesel commercial equipment, gasoline construction and mining equipment, and gasoline commercial equipment were combined to produce the weighting factor for the commercial/residential development map layer. As shown in Table 4, heavy-duty vehicle activity was assigned the highest weighting factor in model scenario 1 due to its high EC contribution, followed by commercial/residential development areas representing heavy-duty construction emissions. Total population density map layers were assigned the highest weighting factor in model scenario 2 to identify areas where DPM emissions are likely to impact highly populated areas.

RESULTS AND DISCUSSION

The results of this analysis are presented as suitability maps indicating areas of high (red) to low (light green) suitability. Low suitability includes areas exhibiting generally unfavorable characteristics for placing monitors to measure DPM. For example, we are interested in placing monitors in locations with high population density; therefore, areas of low population density would be classified as low suitability while areas of high population density would be classified as high suitability. Medium suitability is defined as areas with some suitable features that heighten the importance of an area for emissions activity, population density, or meteorology. High suitability indicates areas where the significant and favorable features (i.e., DPM emissions sources, population density, and wind direction) converge.

Figure 9 shows the results of the hot spot (model scenario 1) suitability analysis without considering meteorological influence. Areas of potentially high DPM emissions are in red; these are areas in which a monitor would be well placed to measure a mix of sources that emit DPM. Areas near the intersection of Interstate 10 and 17, and the downtown section of Interstate 10, are identified as suitable areas for monitoring DPM. Figure 10 shows the model scenario 1 incorporating the influence of meteorology. The existing Bethune School monitoring site is identified by Figure 10 as suitable for monitoring DPM. The area identified as highly suitable surrounding downtown Phoenix shifts to the southwest when meteorology is used in the model. Figure 11 shows the results from the model scenario 2 analysis for total population accounting for meteorology. When population and meteorology are considered in the suitability model, the suitable areas are located just southwest of high population density regions. For example, the

JLG Supersite is located near an area of high population. When the predominately southwesterly wind influence is added to the model, the resulting map indicates that regions just to the southwest of the Supersite are identified as potentially suitable for monitoring population exposure to DPM. Likewise, Guadalupe is identified as a potentially suitable area for monitoring population exposure to DPM.

CONCLUSIONS

The GIS-based Spatial Analyst tool was applied to identify regions in the Phoenix area that are predicted to have high concentrations of DPM. Many sources of geographic data were used to develop the model. Areas of high DPM emissions were identified in the resulting suitability maps, as expected, along main truck routes, in highly industrial regions, and in areas of high construction activity. With the boom of construction in the Phoenix area, areas that have undergone substantial residential and commercial construction appeared on the resulting suitability maps as being likely locations of DPM emissions. Incorporating meteorology into the suitability model significantly affected the results by introducing a southwestern shift in the areas identified as potentially suitable for monitoring DPM.

When the existing air toxics monitoring sites are overlaid on the resulting suitability maps, it appears that the two long-term air toxics monitoring sites located in central Phoenix, Bethune School, and Phoenix Supersite, are well located to monitor a mix of DPM emissions sources. However, when meteorology is accounted for, the areas identified as suitable shift to the southwest, consistent with the predominant southwesterly wind direction in Phoenix. It is important to note that the monitoring objectives for the Phoenix Supersite and Bethune School were not originally set to investigate DPM impacts. When population density is considered, the existing two monitoring sites in central Phoenix are located in areas of high total population density.

This analysis demonstrates the utility and effectiveness of using spatial data with GIS tools to better understand urban-scale emissions patterns, their potential impact on population, and possible locations for placing monitoring sites to measure impacts of DPM. This analysis also demonstrated the importance of considering meteorology.

The results from these analyses should be considered preliminary; analyses were performed to demonstrate the usefulness of the spatial suitability analysis techniques. Several other data types and analyses should be considered in future suitability analyses to enhance results:

- Improved activity information for rail, heavy-duty diesel, and airport diesel emissions,
- Continued assessment and refinement of surrogates for diesel construction,
- Investigation of the relationship between EC and BC data in Phoenix and a comparison of EC and BC data to suitability model results, and
- Investigation of seasonal variability in DPM sources (and meteorology) on the results.

ACKNOWLEDGEMENTS

This work was funded by the Arizona Department of Environmental Quality (ADEQ) to support the ADEQ toxics monitoring program.

REFERENCES

1. U.S. Environmental Protection Agency. *Health assessment document for diesel engine exhaust*; Prepared by U.S. Environmental Protection Agency, Office of Research and Development (National Center for Environmental Assessment), EPA/600/8-90/057F, Washington D.C., 2002.

2. Dockery, D. W.; Pope, C. A., III Acute respiratory effects of particulate air pollution; *Annu Rev Public Health* **1994**, *15*, 107-132.

3. Dockery, D. W.; Pope, C. A.; Xu, X. P.; Spengler, J. D.; Ware, J. H.; Fay, M. E.; Ferris, B. G.; Speizer, F. E. An association between air pollution and mortality in six U.S. cities; *New Engl. J. Med.* **1993**, *329*, 1753-1759.

4. Verma, D. K.; Finkelstein, M. M.; Kurtz, L.; Smolynec, K.; Eyre, S. Diesel exhaust exposure in the Canadian railroad work environment; *Applied Occupational and Environmental Hygiene* **2003**, *18(1)*, 25-34.

5. Schauer, J. J.; Rogge, W. F.; Hildemann, L. M.; Mazurek, M. A.; Cass, G. R.; Simoneit, B. R. T. Source apportionment of airborne particulate matter using organic compounds as tracers; *Atmos. Environ.* **1996**, *30(22)*, 3837-3855.

6. Kim, E.; Hopke, P. K. Improving source identification of fine particles in a rural northeastern U.S. area utilizing temperature-resolved carbon fractions; *J. Geophys. Res.* **2004**, *109(D9)*.

7. Kim, E.; Hopke, P. K.; Edgerton, E. S. Improving source identification of Atlanta aerosol using temperature resolved carbon fractions in positive matrix factorization; *Atmos. Environ.* **2004**, *38*, 3349-3362.

8. Brown, S. G.; Hafner, H. R.; Shields, E. Source apportionment of Detroit air toxics data with positive matrix factorization. Presented at Air & Waste Management Association Symposium on Air Quality Measurement Methods and Technology: Research Triangle Park, NC, 2004

9. Weiland, S. K.; Mundt, K. A.; Ruckman, A.; Keil, U. Self-reported wheezing and allergic rhinitis in children and traffic density on street of residence; *Annals of Epidemiology* **1994**, *4*, 243-247.

10. Ciccone, G.; Forastiere, F.; Agabiti, N.; Biggeri, A.; Bisanti, L.; Chellini, E.; Corbo, G.; Dell'Orco, V.; Dalmasso, P.; Volante, T. F.; Galassi, C.; Piffer, S.; Renzoni, E.; Rusconi, F.; Sestini, P.; Viegi, G. Road traffic and adverse respiratory effects in children; *Occupational & Environmental Medicine* **1998**, *55*, 771-778.

11. Zhu, Y.; Hinds, W. C.; Kim, S.; Sioutas, C. Concentration and size distribution of ultrafine particles near a major highway; *J. Air & Waste Manag. Assoc.* **2002**, *52*, 1032-1042.

12. Lim, H.-J.; Turpin, B. J.; Edgerton, E.; Hering, S. V.; Allen, G. A.; Maring, H.; Solomon, P. Semicontinuous aerosol carbon measurements: comparison of Atlanta supersite measurements; *J. Geophys. Res.* **2003**, *108*(*D7)*, 8419.

13. Watson, J. G.; Chow, J. C. Comparison and evaluation of in situ and filter carbon measurements at the Fresno Supersite; *J. Geophys. Res.* **2002**, *107*(*D21)*.

14. Hafner, H. R.; Brown, S. G.; Penfold, B. M.; Funk, T. H. Use of GIS tool to investigate suitable diesel emissions monitoring locations in Detroit, Michigan. Presented at Air & Waste Management Association Symposium on Air Quality Measurement Methods and Technology: Research Triangle Park, NC, 2004

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

16. Perez, O. M.; Telfer, T. C.; Beveridge, M. C. M.; Ross, L. G. Geographical Information Systems (GIS) as a simple tool to aid modelling of particulate waste distribution at marine fish cage sites; *Estuarine, Coastal, and Shelf Science* **2002**, *54*, 761-768.

17. Matejicek, L.; Benesova, L.; Tonika, J. Ecological modelling of nitrate pollution in small river basins by spreadsheets and GIS; *Ecological Modelling* **2003**, *170*, 245-263.

18. Basnyat, P.; Teeter, L. D.; Lockaby, B. G.; Flynn, K. M. The use of remote sensing and GIS in watershed level analyses of non-point source pollution problems; *Forest Ecology and Management* **2000**, *128*, 65-73.

19. *SPECIATE database, version 3.2*. U.S. Environmental Protection Agency, 2002. Available at: <http://www.epa.gov/ttn/chief/software/speciate/index.html>.

20. Gillies, J. A.; Gertler, A. W. Comparison and evaluation of chemically speciated mobile source PM2.5 particulate matter profiles; *J. Air & Waste Manage. Assoc.* **2000**, *50*, 1459-1480.

21. Schauer, J. J.; Cass, G. R. Source apportionment of wintertime gas-phase and particle-phase air pollutants using organic compounds as tracers; *Environ. Sci. Technol.* **2000**, *34*(*9)*, 1821-1832.

22. Schauer, J. J.; Kleeman, M. J.; Cass, G. R.; Simoneit, B. R. T. Measurement of emissions from air pollution sources. 2. C_1 through C_{30} organic compounds from medium duty diesel trucks; *Environ. Sci. Technol.* **1999**, *33*, 1578-1587.

23. Main, H. H.; Funk, T. H.; Penfold, B. M. *Arizona Department of Environmental Quality toxics network assessment*; STI-901483-2226-FR; Report prepared for the Arizona Department of Environmental Quality by Sonoma Technology, Inc., Petaluma, CA, 2002.

TABLES AND FIGURES

Table 1. Major source categories, emissions, and percent contribution of total $PM_{2.5}$ for Maricopa and Pinal counties as reported in the 1999 NEI. Note these sources combined account for 90% of total $PM_{2.5}$ emissions.

Source Type	**$PM_{2.5}$ (tons/yr)**	**Percentage of Total $PM_{2.5}$ Inventory**
Road Construction (dust and exhaust)	7,036	18.6%
Open Burning	6,959	18.4%
Heavy Construction (dust and exhaust)	3,575	9.4%
All Paved Roads Fugitives (dust)	3,389	8.9%
All Unpaved Roads Fugitives (dust)	2,559	6.7%
Agriculture - Crops Tilling (dust)	2,514	6.6%
Forest Wildfires	1,970	5.2%
Managed Burning, Prescribed	1,499	4.0%
Heavy-Heavy Duty Diesel Vehicles	1,083	2.9%
General Building Construction (dust and exhaust)	1,036	2.7%
Diesel Construction and Mining Equipment	1,012	2.7%
Residential Wood	694	1.8%
Gasoline Lawn and Garden Equipment	416	1.1%

Table 2. Elemental carbon contributions for emissions sources in the $PM_{2.5}$ emission inventory for Maricopa and Pinal counties. The EC weight percent data are from ARB's Speciation Profile Library.

Emissions Source	$PM_{2.5}$ (tons/year)	EC weight percent[a]	EC (tons/year)
All heavy-duty diesel vehicles	1,496	0.264	395
Diesel construction and mining equipment	1,012	0.150	152
All highway vehicles - gasoline	544	0.200	109
Gasoline lawn and garden equipment	416	0.200	83
$PM_{2.5}$ point sources	694	Differs by source type	54
Diesel lawn and garden equipment	127	0.150	19
Diesel commercial equipment	96	0.150	14
Railroad equipment	117	0.100	12
Gasoline construction and mining equipment	28	0.200	6
Aircraft	25	0.150	4
Gasoline commercial equipment	21	0.200	4

[a] California Air Resources Board - Speciation Profiles and Size Fractions (http://www.arb.ca.gov/ei/speciate/speciate.htm)

Table 3. Spatial data assigned to each emissions source.

Emissions Source	Map Layer Assigned to Represent Spatial Distribution of Emissions
Highway vehicles - diesel	Locations of roadways and corresponding heavy-duty vehicle AADT[a], locations of major transportation hubs[b]
Diesel construction and mining equipment	Maps of residential and commercial development areas[c]
Highway vehicles - gasoline	Locations of roadways and corresponding light-duty vehicle AADT[a]
Gasoline lawn and garden equipment	Land use[d]
Diesel lawn and garden equipment	Land use[d]
Railroad equipment	Locations of transportation hubs, locations of railroad links[e]
Diesel commercial equipment	Maps of development areas[d]
Gasoline construction and mining equipment	Maps of development areas[d]
Aircraft	Airport locations[f]
Gasoline commercial equipment	Maps of development areas[d]
$PM_{2.5}$ point sources	Locations of point sources[g]

[a] Arizona Department of Transportation, 2000
[b] National Transportation Atlas Data, Bureau of Transportation Statistics, 2002
[c] Maricopa Association of Governments (MAG), Residential/Commercial Completions, 2000-2003
[d] Maricopa Association of Governments, Existing Land use, 2000
[e] National Transportation Atlas Data, Bureau of Transportation Statistics, 2002
[f] National Transportation Atlas Data, Bureau of Transportation Statistics, 2002
[g] EPA's Aerometric Information Retrieval System (AIRS), now Air Quality System (AQS), 1999

Table 4. Weighting scheme for suitability analysis for monitoring diesel emissions and population exposure to diesel emissions.

Layer	Weighting Scheme (1) Hot Spot	Weighting Scheme (2) Total Population	Weighting Criteria
Density of total population	-	40%	High population density = more suitable
Heavy-duty vehicle activity	20%	12%	High traffic density = more suitable
Light-duty vehicle activity	15%	9%	High traffic density = more suitable
Transportation distribution facility	20%	12%	Close to facility = more suitable
Lawn/garden activity areas	12%	7.2%	High activity density = more suitable
Commercial/residential construction activity areas	20%	12%	High activity density = more suitable
Distance to airports	2%	1.2%	Close to airport = more suitable
Distance to railroads	2%	1.2%	Close to railroad = more suitable
$PM_{2.5}$ point source activity	9%	5.4%	High $PM_{2.5}$ emissions density = less suitable

Figure 1. Map of the Phoenix area depicting long-term air toxics monitoring sites (blue triangles), topography, urban features, and tribal lands (red polygons).

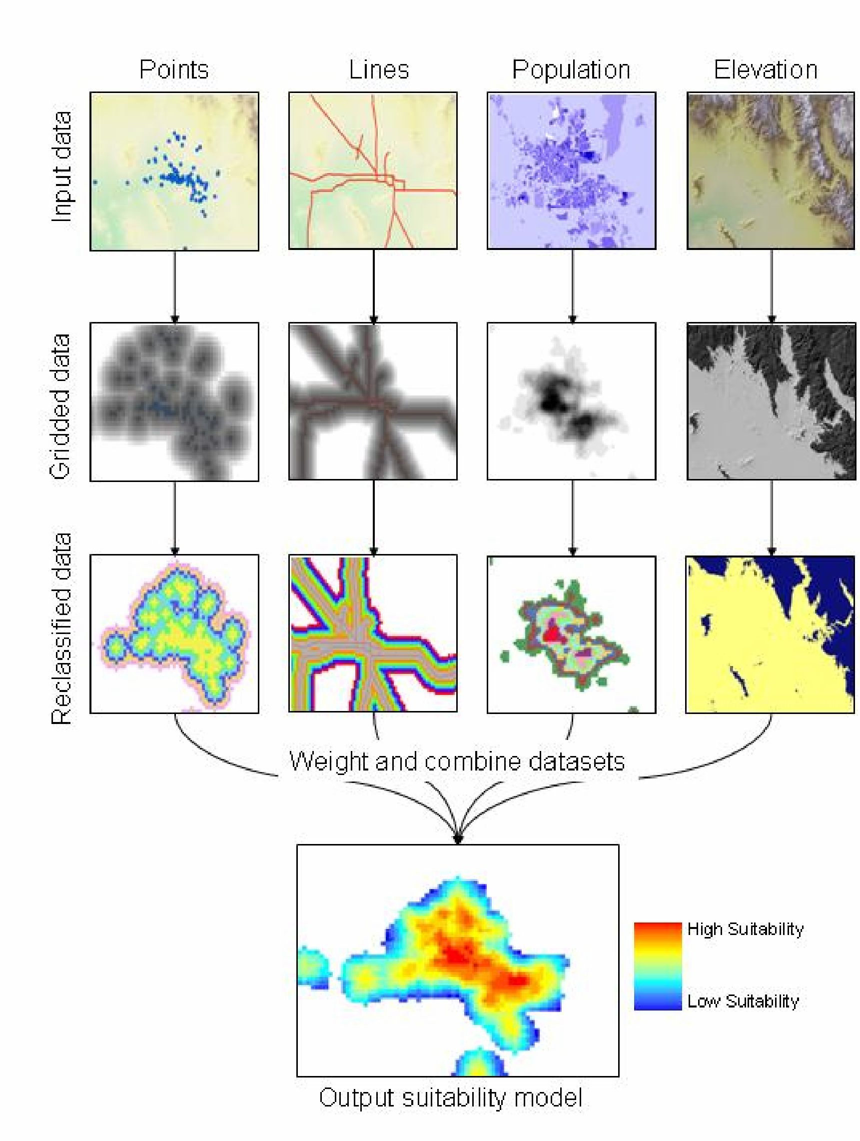


Figure 2. Conceptual approach for building a suitability model.

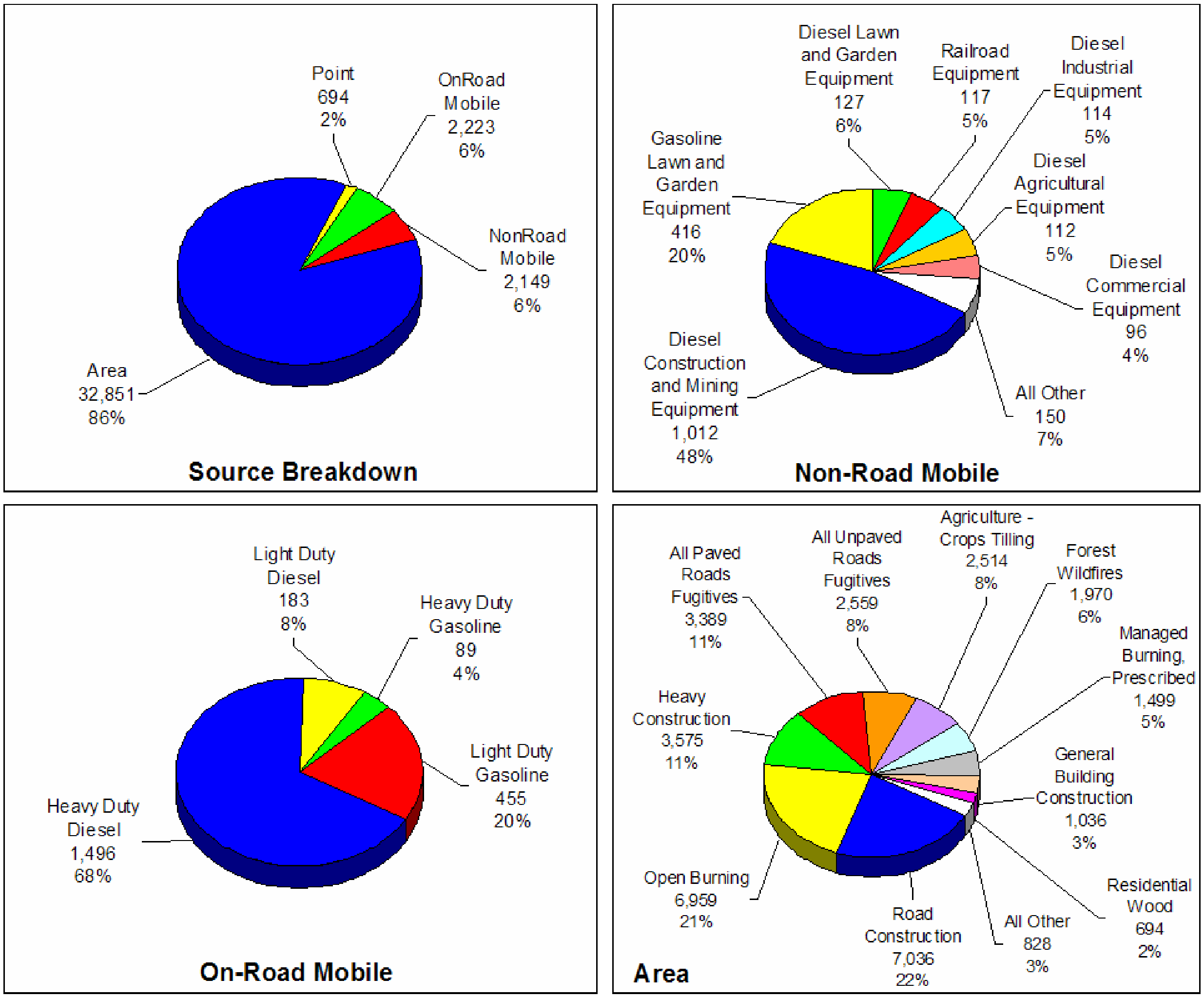
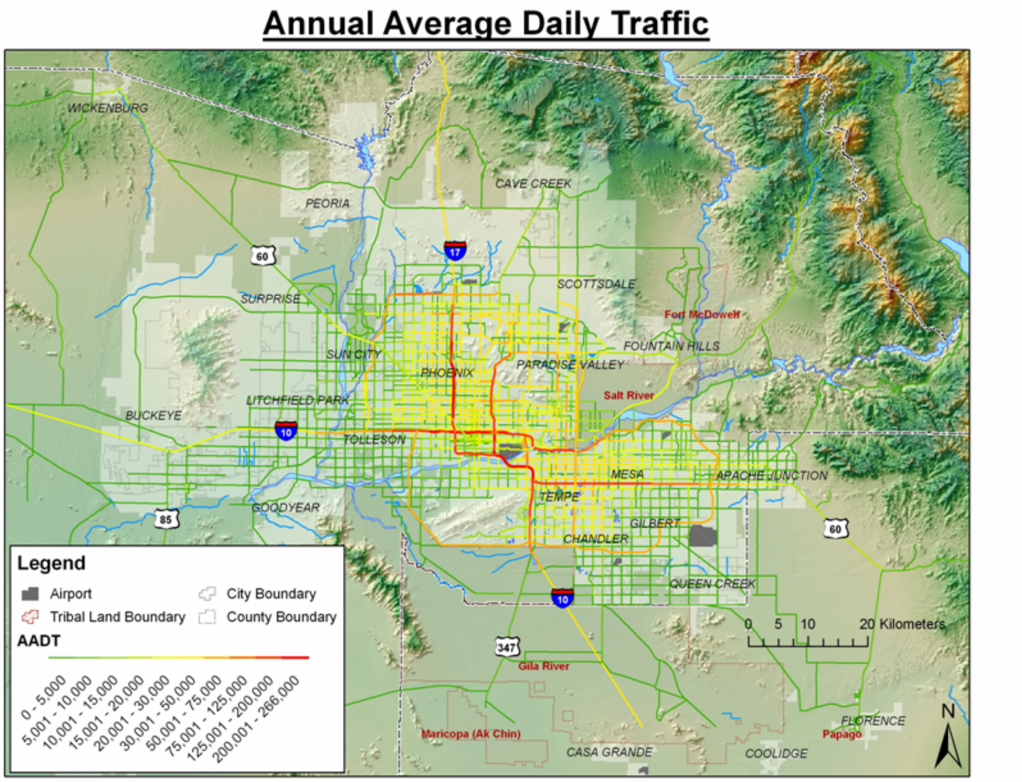


Figure 3. Emissions source contributions to total $PM_{2.5}$ for Maricopa and Pinal counties as reported in the 1999 NEI.

(a)



(b)

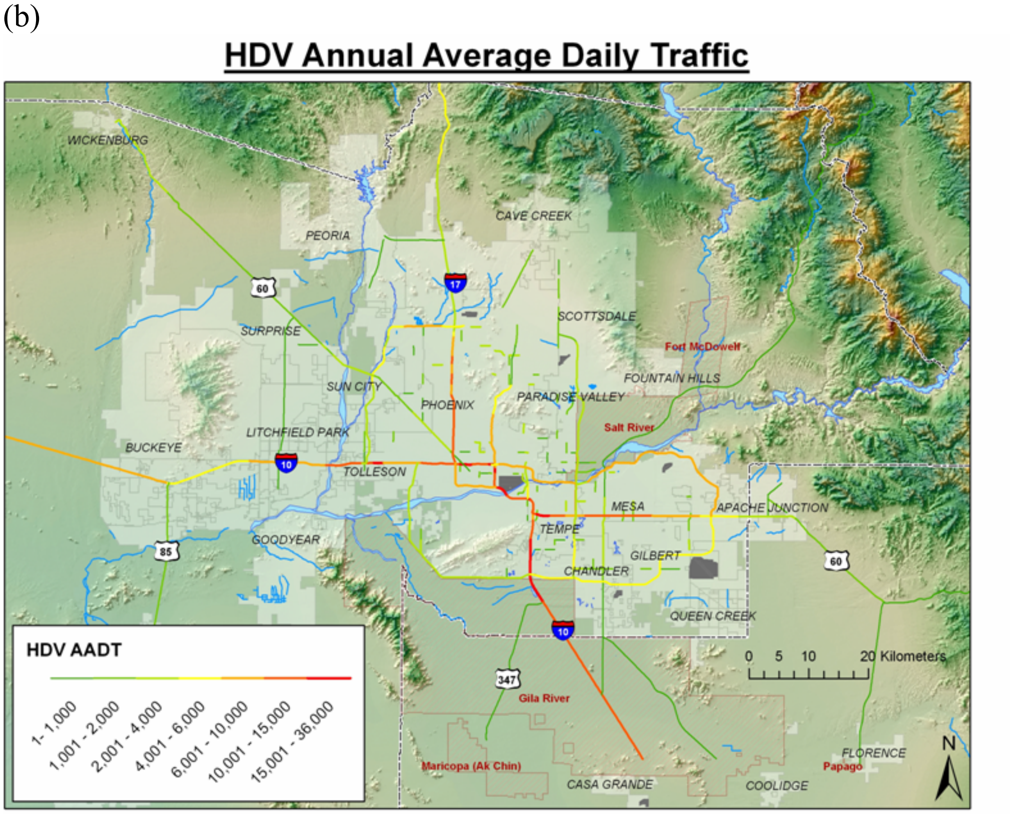


Figure 4. Annual average daily traffic volume (AADT) (a) and heavy-duty vehicle traffic volume (b) for the Phoenix area.

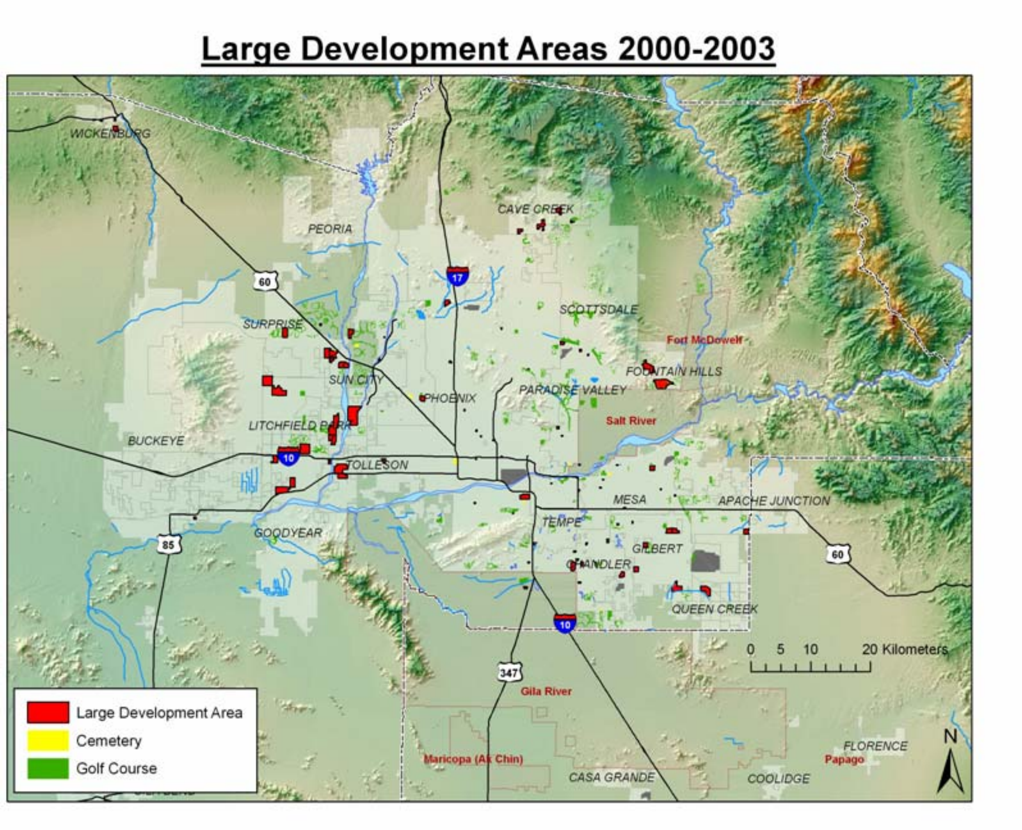


Figure 5. Large development areas, golf courses, and cemeteries in the Phoenix area.

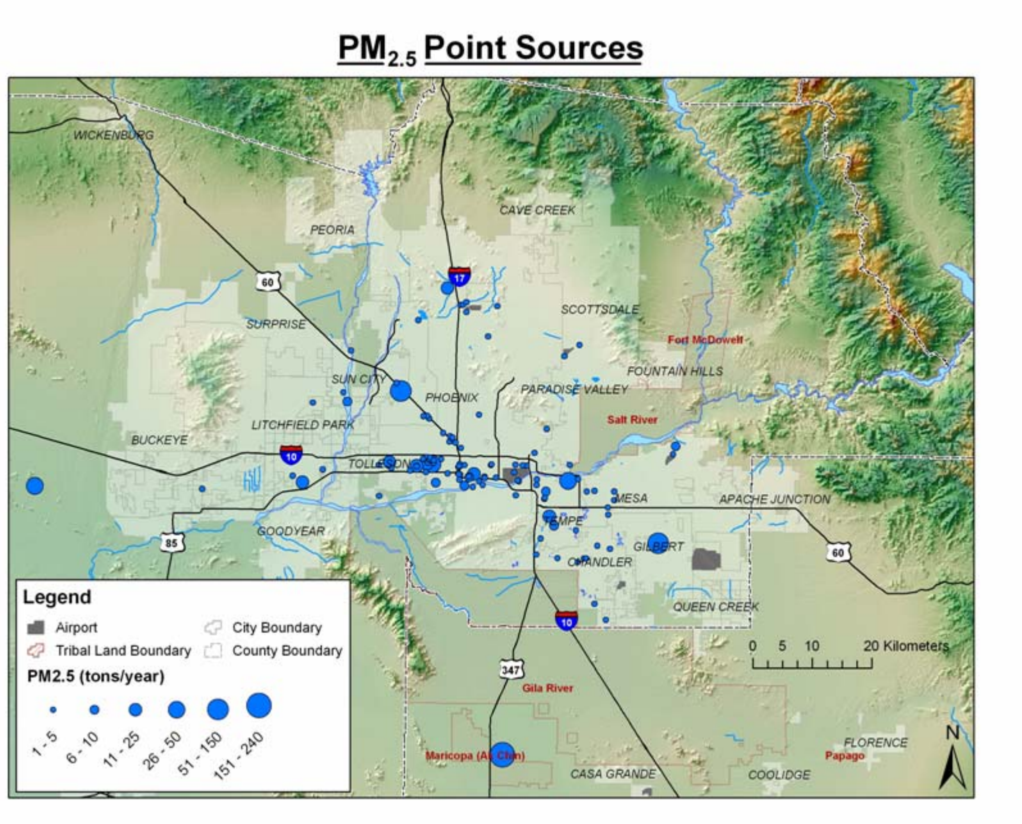


Figure 6. Location and magnitude of emissions for $PM_{2.5}$ point sources in Phoenix.

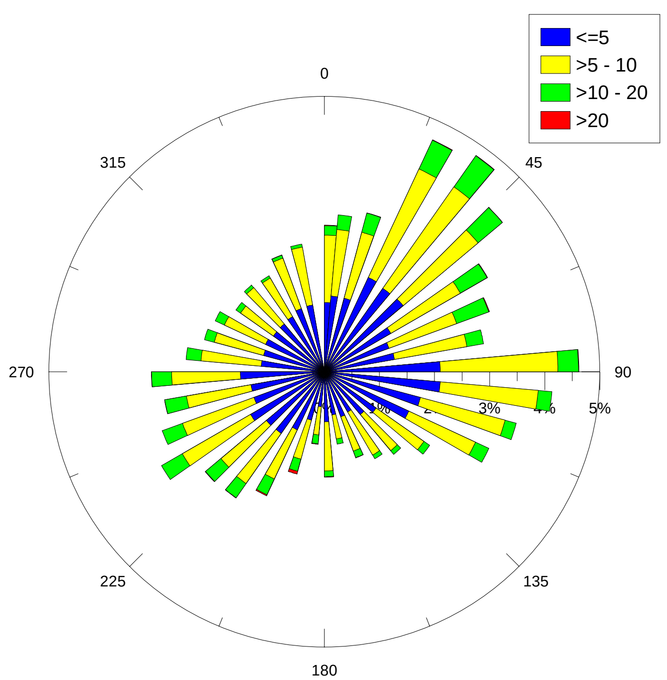


Figure 7. Predominant annual (2001-2003) wind rose for the Phoenix area. Wind speed measured in meters per second.

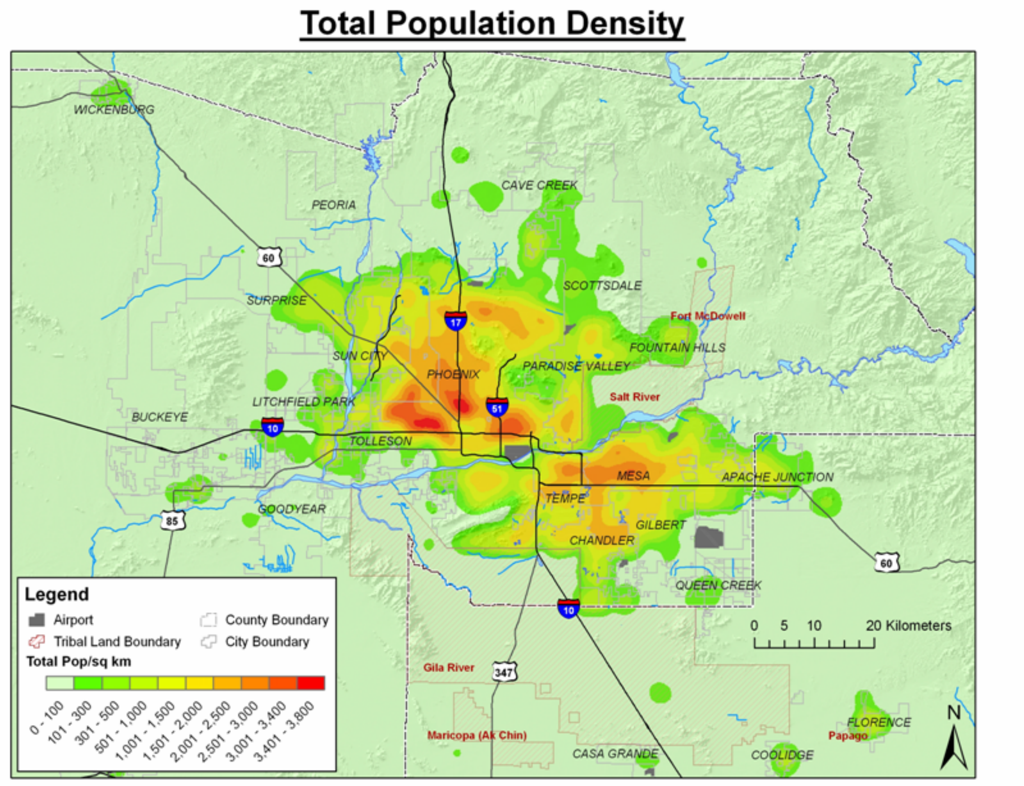


Figure 8. Total population density in the Phoenix area.

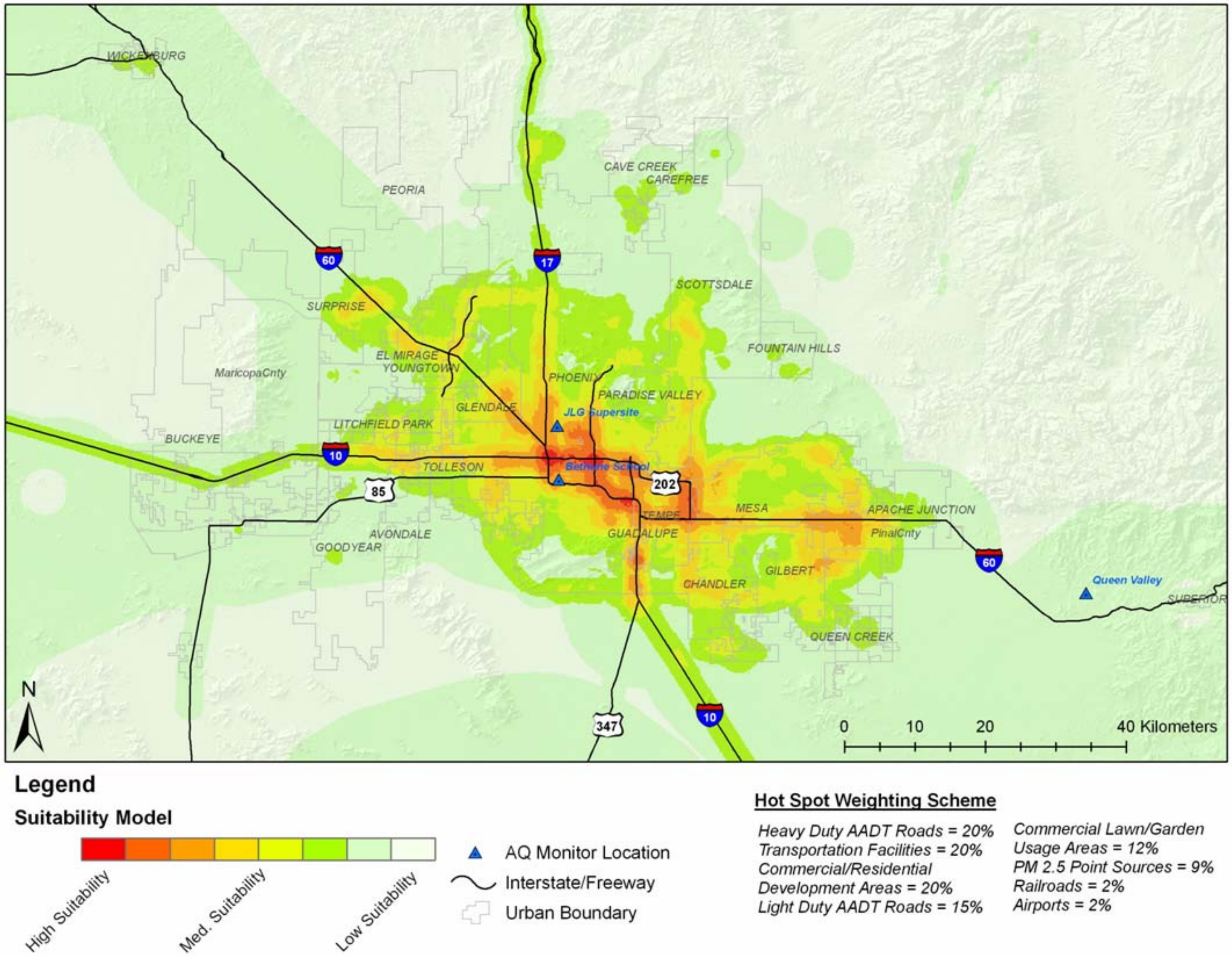


Figure 9. Hot spot suitability analysis without meteorological influence.

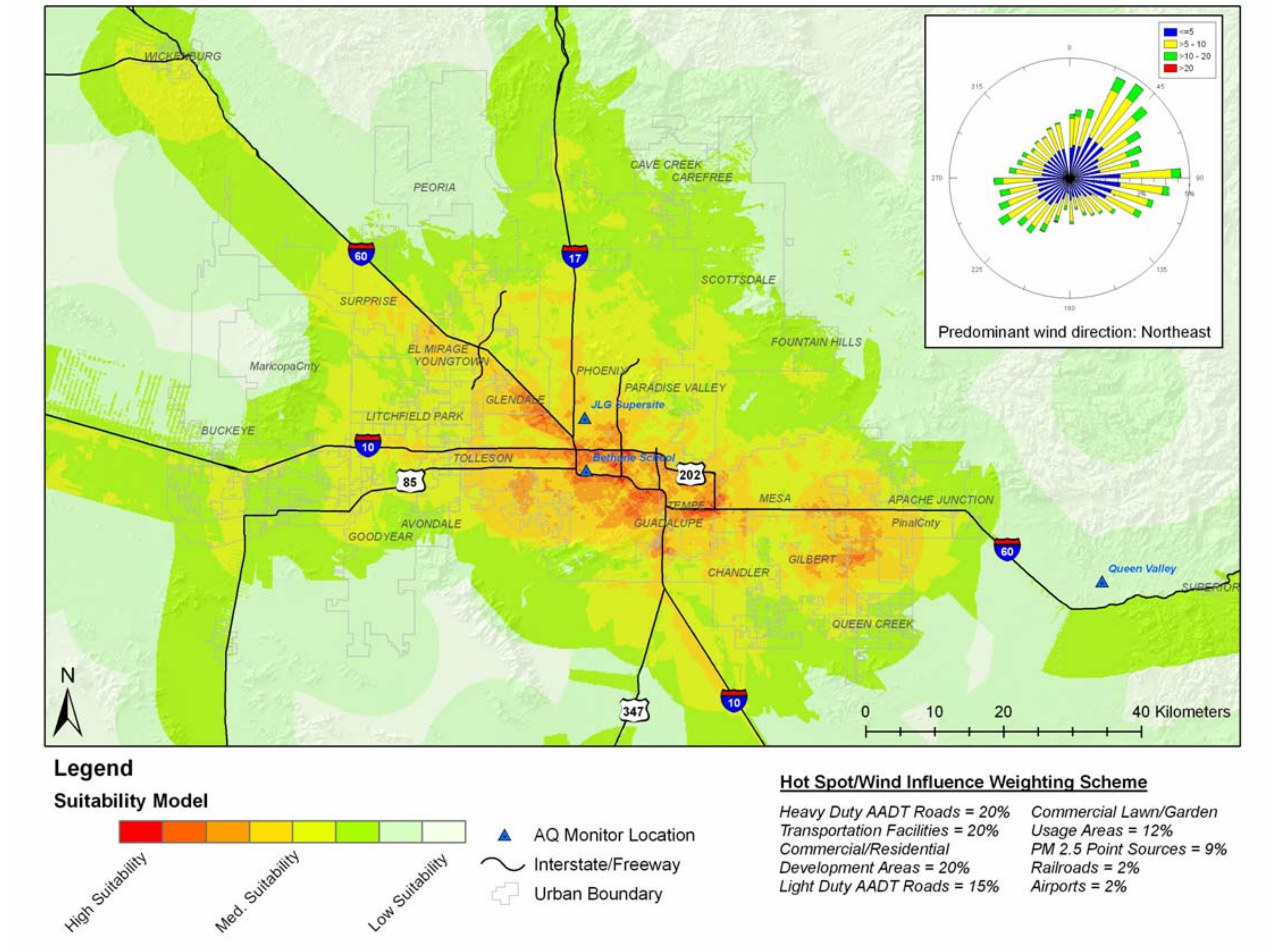


Figure 10. Hot spot suitability analysis with meteorological influence.

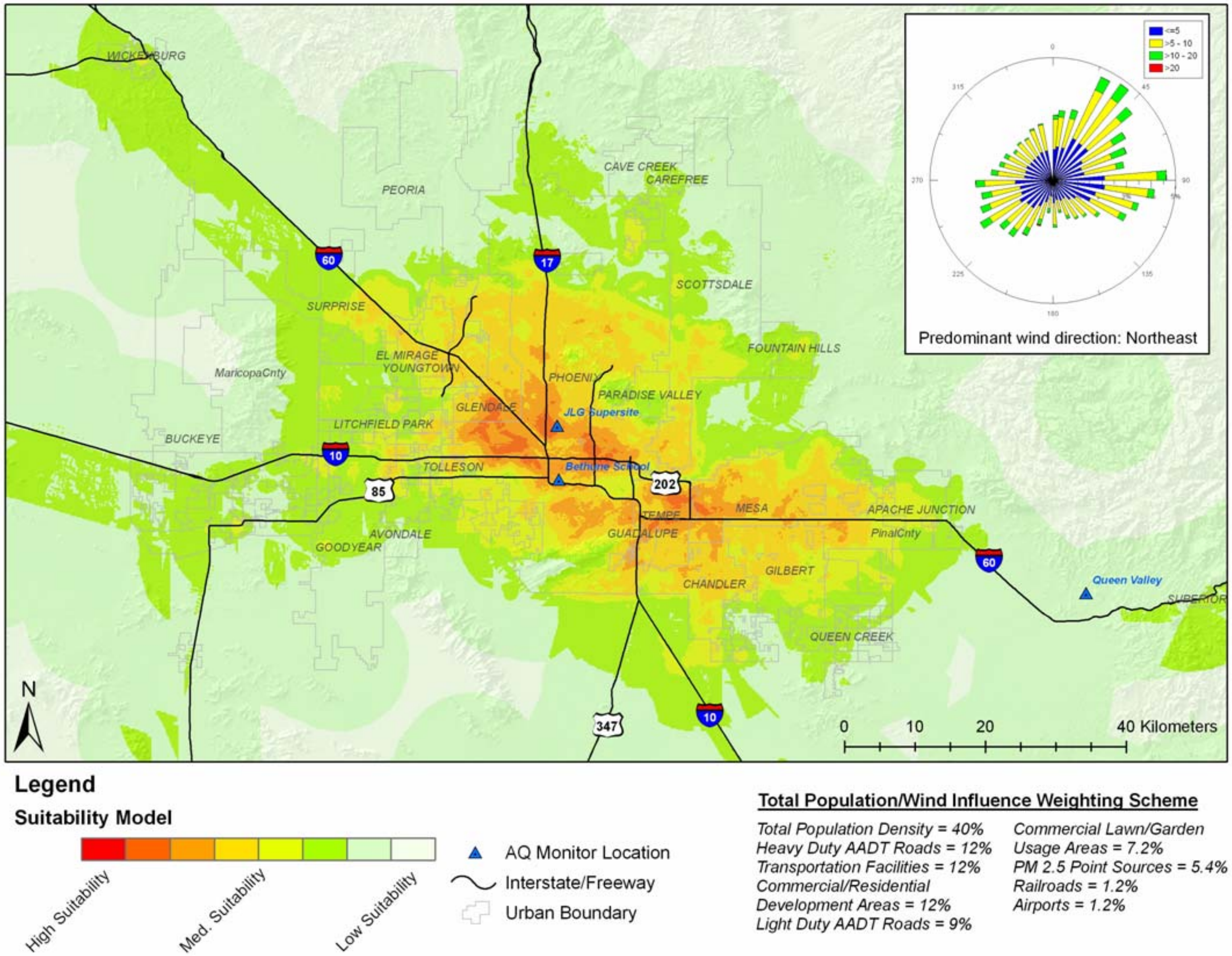


Figure 11. Total population suitability analysis accounting for meteorology.