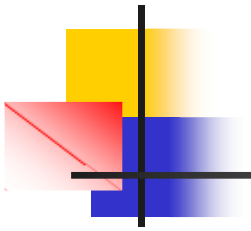
Analytical Techniques Used for Regional-Scale Network Assessments

Prepared by:
Bryan M. Penfold
Hilary R. Hafner
Sean M. Raffuse
Sonoma Technology, Inc.
Petaluma, CA

Kevin Cavender
U.S. EPA OAQPS

Presented to:
2006 National Air Monitoring Conference
Las Vegas, NV
November 6-9, 2006



Overview

Goal:

Provide examples and guidance on analytical techniques that can be used to evaluate a network's effectiveness and efficiency relative to its objectives and costs

Agenda:

- Thought process for a network assessment
- Analysis tools and resources
- "Simpler" analysis examples
- More complex analysis examples

- What are some reasons why an assessment might be needed?
- What questions are we trying to answer in the assessment?

Network Assessment Analytical Techniques

Analyses can be used to

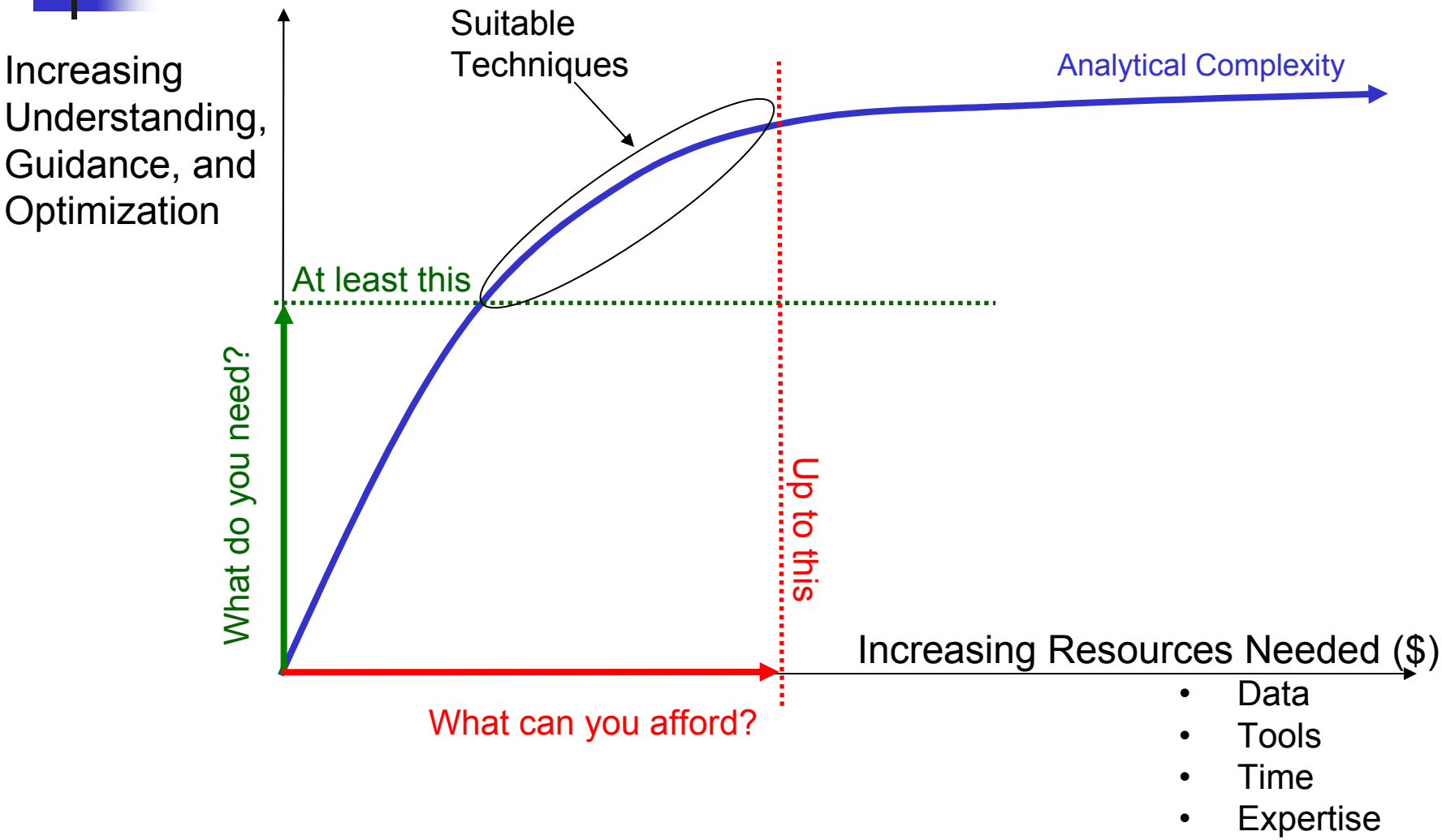
- Identify potential redundancies or to determine the adequacy of existing monitoring sites
- Identify potential adjustments to protect today's population
- Address multiple, interrelated air quality issues
- Maintain the ability to understand long-term historical air quality trends
- Refocus resources on pollutants that are new or persistent challenges and deemphasize monitoring for pollutants that are better understood

Monitoring Networks Support Many Objectives

- Meet national compliance requirements
- Evaluate air quality models
- Evaluate emission inventories
- Support source apportionment
- Understand temporal variability
- Track long-term trends
- Monitor specific sources
- Monitor areas of maximum precursor or primary emissions
- Monitor the background concentration
- Characterize transport
- Support interpolation and mapping
- Assist forecasting
- Public reporting (AQI)

What others can you think of?

Network Assessment Techniques: One Size Does Not Fit All



Analysis Techniques – Broadly

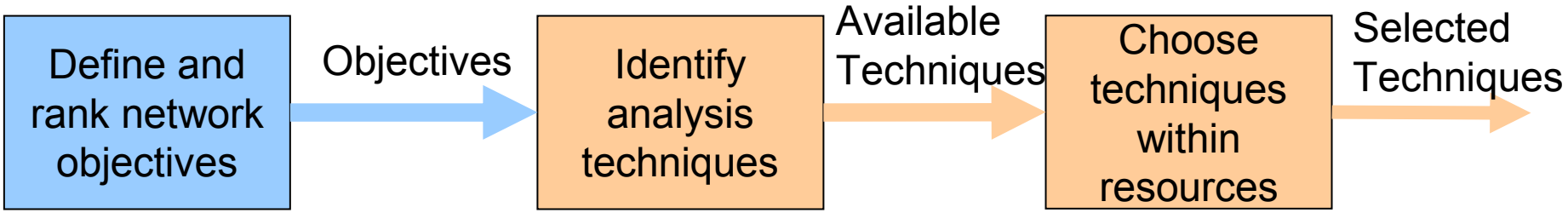
Site-by-site	Bottom-up	Network optimization
What's the relative value of current sites?	Where are there deficiencies in the network?	What is the "best" network design?

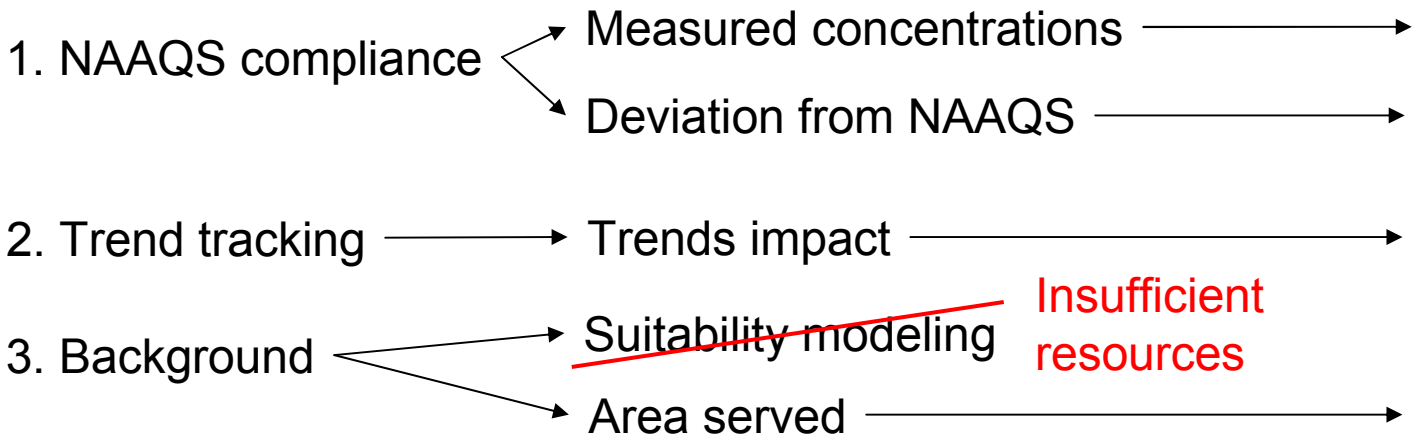
Increasing complexity (in general)

- A rigorous network assessment will typically have to incorporate both site-by-site and bottom-up analysis techniques.
- Network optimization entails analyzing hypothetical network scenarios.

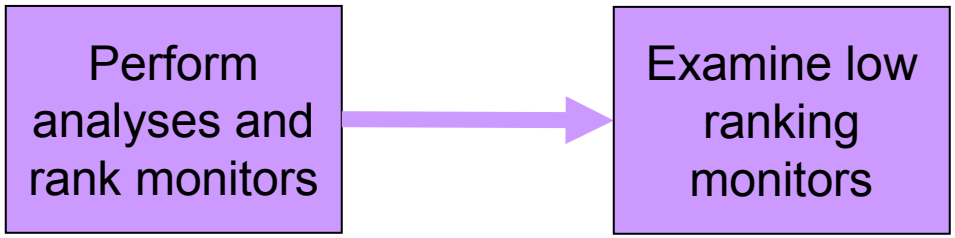
What Is the Relative Value of Current Ozone Sites? (1 of 2)

An example site-by-site analysis flow





What Is the Relative Value of Current Ozone Sites? (2 of 2)



	NAAQS	Deviation	Trends	Area	Overall
Site 1	1	1	2	3	7
Site 2	3	2	1	1	7
Site 3	2	3	3	2	10

Low ranking monitors should be examined carefully and case-by-case

- What was the original monitor objective?
- Is this monitor fulfilling secondary objectives?
- Possible reallocation of resources: locations, pollutants, technologies

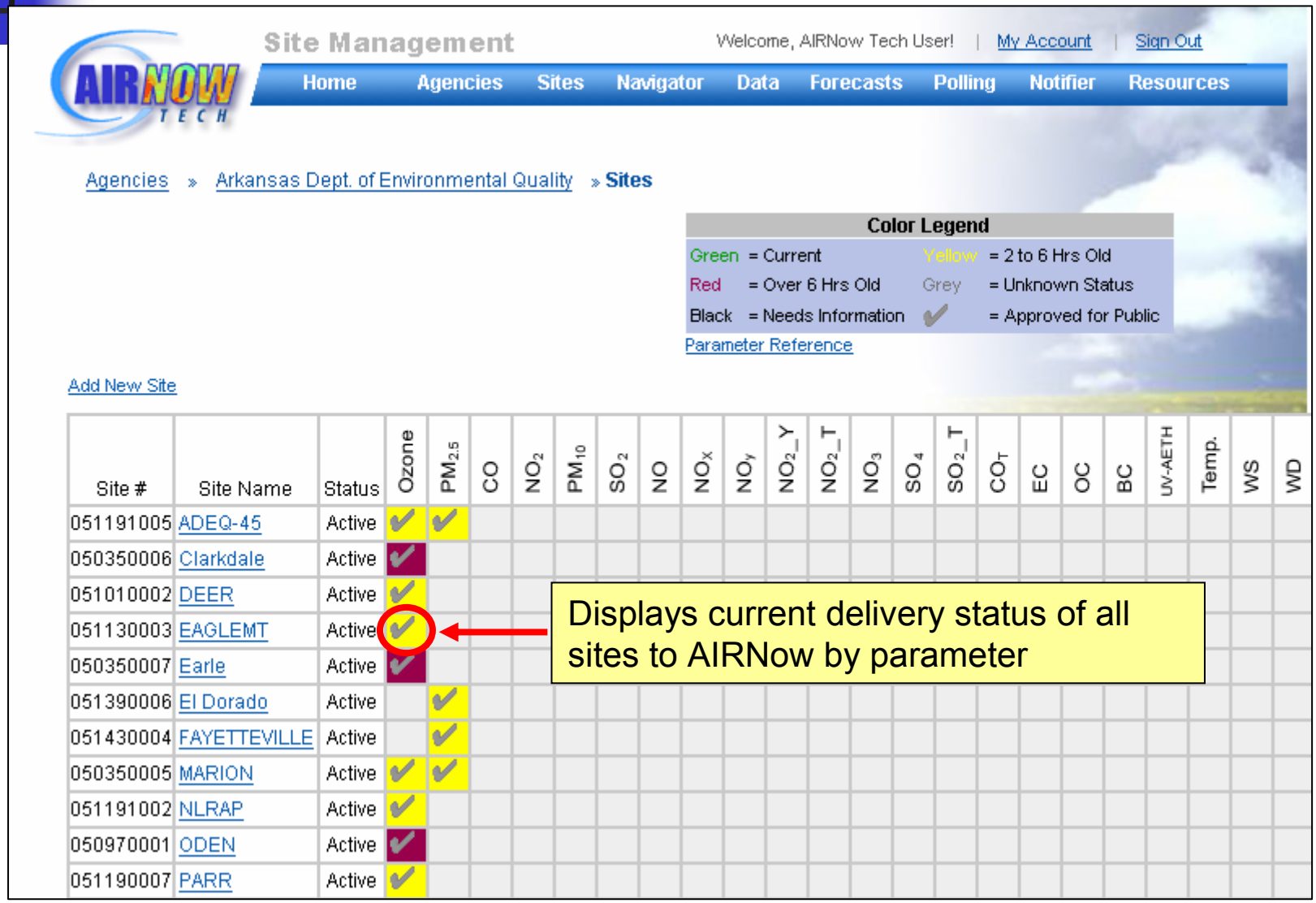
Analysis Tools (1 of 2)

What resources are useful for network assessment?

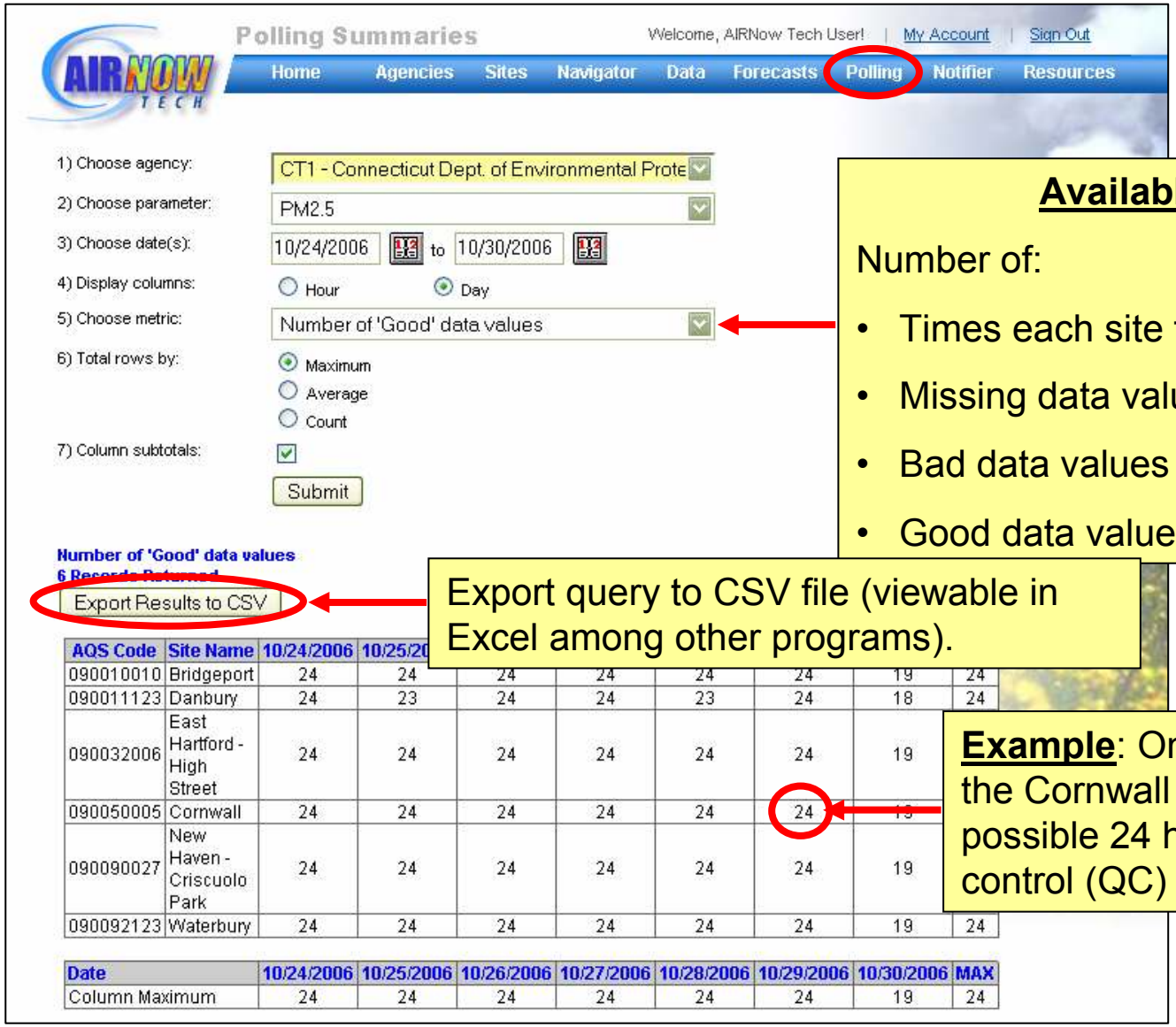
Data Sources

- **EPA AirData** web page gives you access to yearly summaries of U.S. air pollution data, taken from EPA's air pollution databases. Types of data include emissions and monitoring.
- **EPA AIRNow Tech** web page gives you access to AIRNow observational data. Within AIRNow-Tech are the Navigator and Data tools. The Navigator tool is a customizable, air quality GIS tool that allows you to display site information with multiple geographic, pollutant, and meteorological features. The Data tool allows you to create personalized site lists, access predefined queries, and download AIRNow observational data.

Site Management



Polling Summaries



AQS Code	Site Name	10/24/2006	10/25/20						
090010010	Bridgeport	24	24	24	24	24	24	19	24
090011123	Danbury	24	23	24	24	23	24	18	24
090032006	East Hartford - High Street	24	24	24	24	24	24	19	
090050005	Cornwall	24	24	24	24	24	24		
090090027	New Haven - Criscuolo Park	24	24	24	24	24	24	19	
090092123	Waterbury	24	24	24	24	24	24	19	24

Date	10/24/2006	10/25/2006	10/26/2006	10/27/2006	10/28/2006	10/29/2006	10/30/2006	MAX
Column Maximum	24	24	24	24	24	24	19	24

Available Metrics

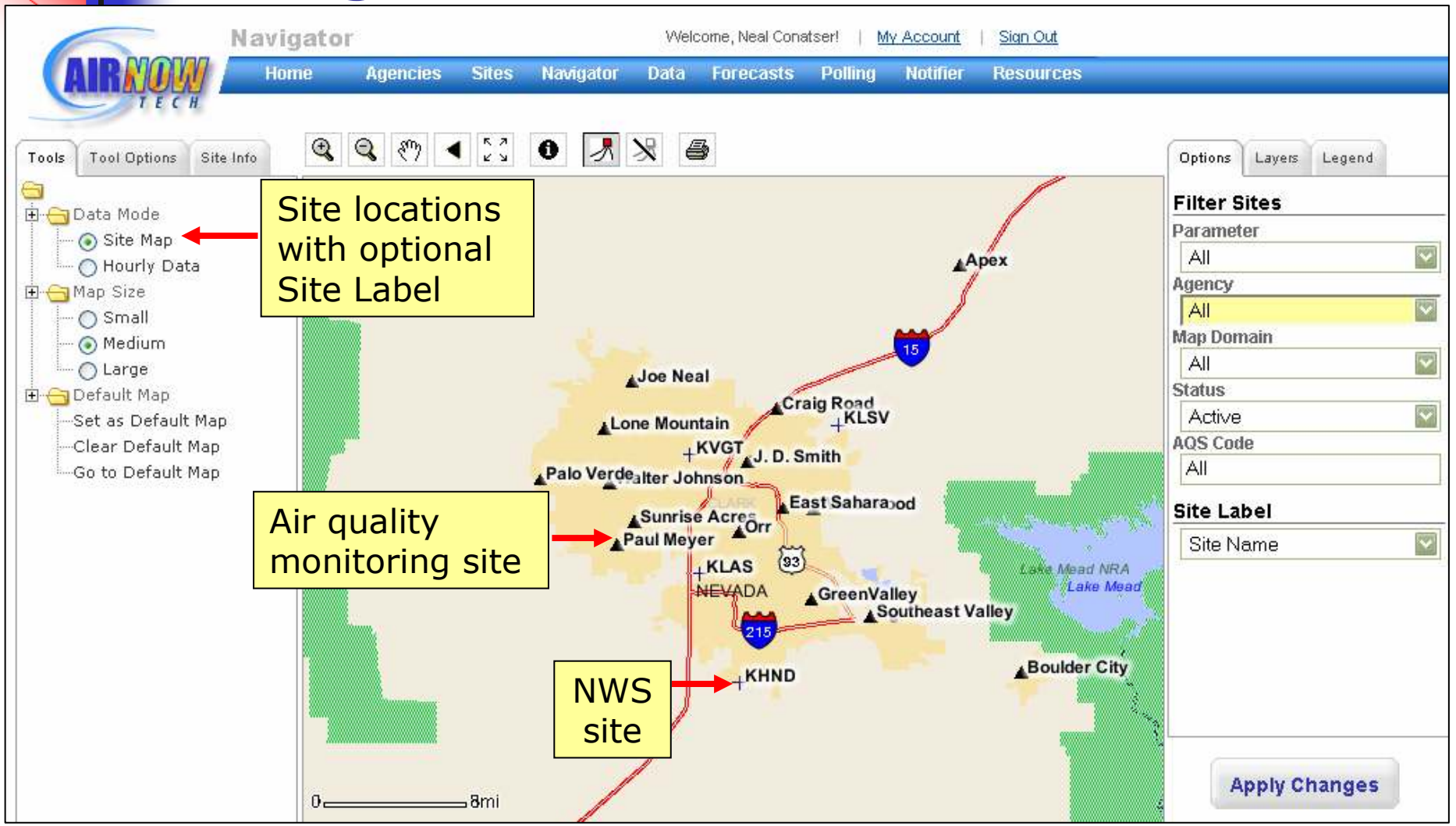
Number of:

- Times each site failed QC
- Missing data values
- Bad data values
- Good data values

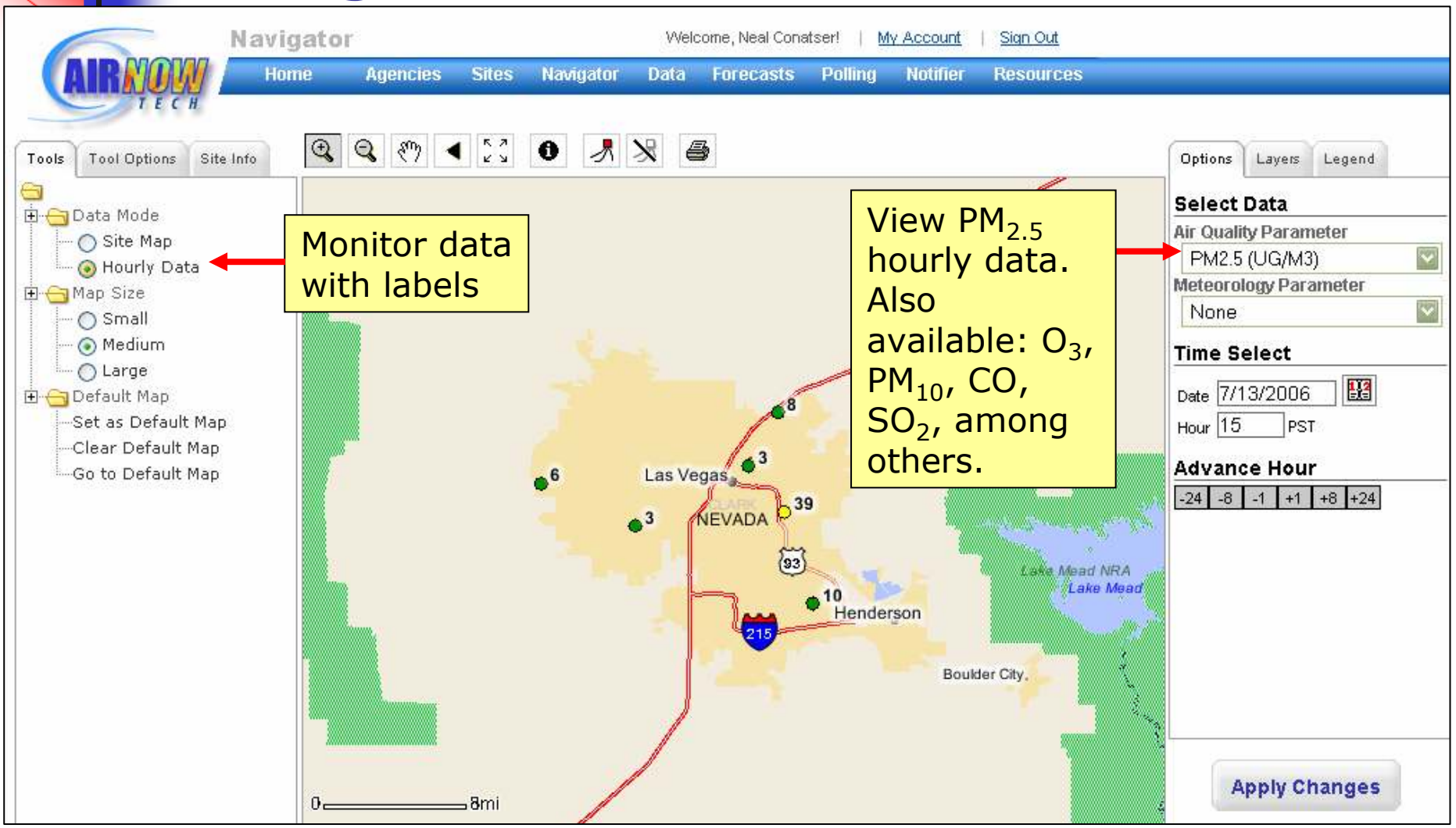
Export query to CSV file (viewable in Excel among other programs).

Example: On 10/29/06, $PM_{2.5}$ data from the Cornwall site met 24 out of a possible 24 hours of specified quality control (QC) criteria.

Navigator GIS



Navigator GIS



Analysis Tools (2 of 2)

What resources are useful for network assessment?

Tools

- **Geographic Information Systems (GIS)** are systems for management, analysis, and display of geographic information. GIS software is available for purchase from ESRI, MapInfo, AutoDesk, etc.
- **Statistical software, database packages** such as Microsoft Excel, Microsoft Access, Grapher, SAS, SYSTAT, etc. These software packages allow you to organize, manipulate, create, analyze, and display data.

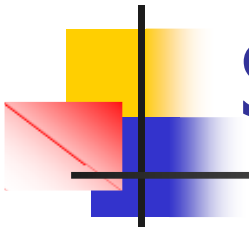
Useful Web Links

Where can I access the resources useful for network assessment?

- EPA AirData:
  http://www.epa.gov/air/data
- EPA AIRNow Tech:
  http://www.airnowtech.org
- Geographic Information Systems (GIS):
  http://www.esri.com
  http://www.mapinfo.com
  http://usa.autodesk.com
- Statistical software, database packages:
  http://www.systat.com
  http://www.goldensoftware.com/products/grapher/graph.shtml
  http://www.sas.com/software

Methods for Technical Assessment

*	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



Site-by-site Analysis Techniques

- Assign a ranking to individual monitors based on a particular metric.
- Are good for assessing which monitors might be candidates for modification or removal.
- 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.

Technique	Complexity	Objectives Assessed
Number of other parameters monitored at the site	*	Overall site value Model evaluation Source apportionment
Trends impact	* to **	Trend tracking Historical consistency Emission reduction evaluation
Measured concentrations	**	Maximum concentration location Model evaluation Regulatory compliance Population exposure
Deviation from NAAQS	**	Regulatory compliance Forecasting assistance
Area served	**	Spatial coverage Interpolation Background concentration

Technique	Complexity	Objectives Assessed
Monitor-to-monitor correlation	** to ***	Model evaluation Spatial coverage Interpolation
Population served	***	Population exposure Environmental justice
Principal component analysis	***	Background concentration Forecasting assistance
Removal bias	***	Regulatory compliance Model evaluation Spatial coverage Background concentration Interpolation

Bottom-up Analyses

- Examine the phenomena that are thought to cause high pollutant concentrations and/or population exposure, such as emissions, meteorology, and population density.
- 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.
- Can be complex and require significant resources (time, data, tools, and analytical skill).

Site-by-site and bottom-up analyses are best performed in combination. Site-by-site analyses typically identify network redundancies while bottom-up analyses identify network "holes" or deficiencies.

Technique	Complexity	Objectives Assessed
Emission Inventory	** to ****	Emission reduction evaluation Maximum precursor location
Population density	**	Population exposure Environmental justice
Population change	***	Population exposure Environmental justice Maximum precursor location
Suitability modeling	****	Population exposure Environmental justice Source-oriented Model evaluation Maximum concentration location Background concentration Transport/border characterization
Photochemical modeling	****	Maximum concentration location Source-oriented Transport/border characterization Population exposure Background concentration

Network Optimization Methods

- Are a holistic approach to examining an air monitoring network.
- Typically assign scores to different network scenarios; alternative network designs can be compared with the current (base-case) design.

Technique	Complexity	Objectives Assessed
Monitor-to-monitor correlation	** to ***	Model evaluation Spatial coverage Interpolation
Principal Component Analysis	***	Background concentration Forecasting assistance
Removal bias	***	Regulatory compliance Model evaluation Spatial coverage Background concentration Interpolation
Positive matrix factorization	****	Source apportionment Emission inventory evaluation

Network Assessment Analysis Examples

Easier to do

More complex

- Parameters monitored
- Regional/local versus national comparison
- Trends impacts
- Measured concentrations
- Deviation from NAAQS
- Monitor-to-monitor correlation
- Emission inventory (county-level and gridded)
- Population change
- Population served
- Area served
- Removal bias
- Suitability modeling
- Principle component analysis
- Positive matrix factorization

Parameters Monitored

- Motivation:
  - Monitors that are collocated with other measurements at a particular site are more valuable than sites that measure fewer parameters.
  - Operating costs can be leveraged among several instruments at these sites
- Resources needed:
  - Monitor information from the Air Quality System (AQS)
  - Site histories from annual reports

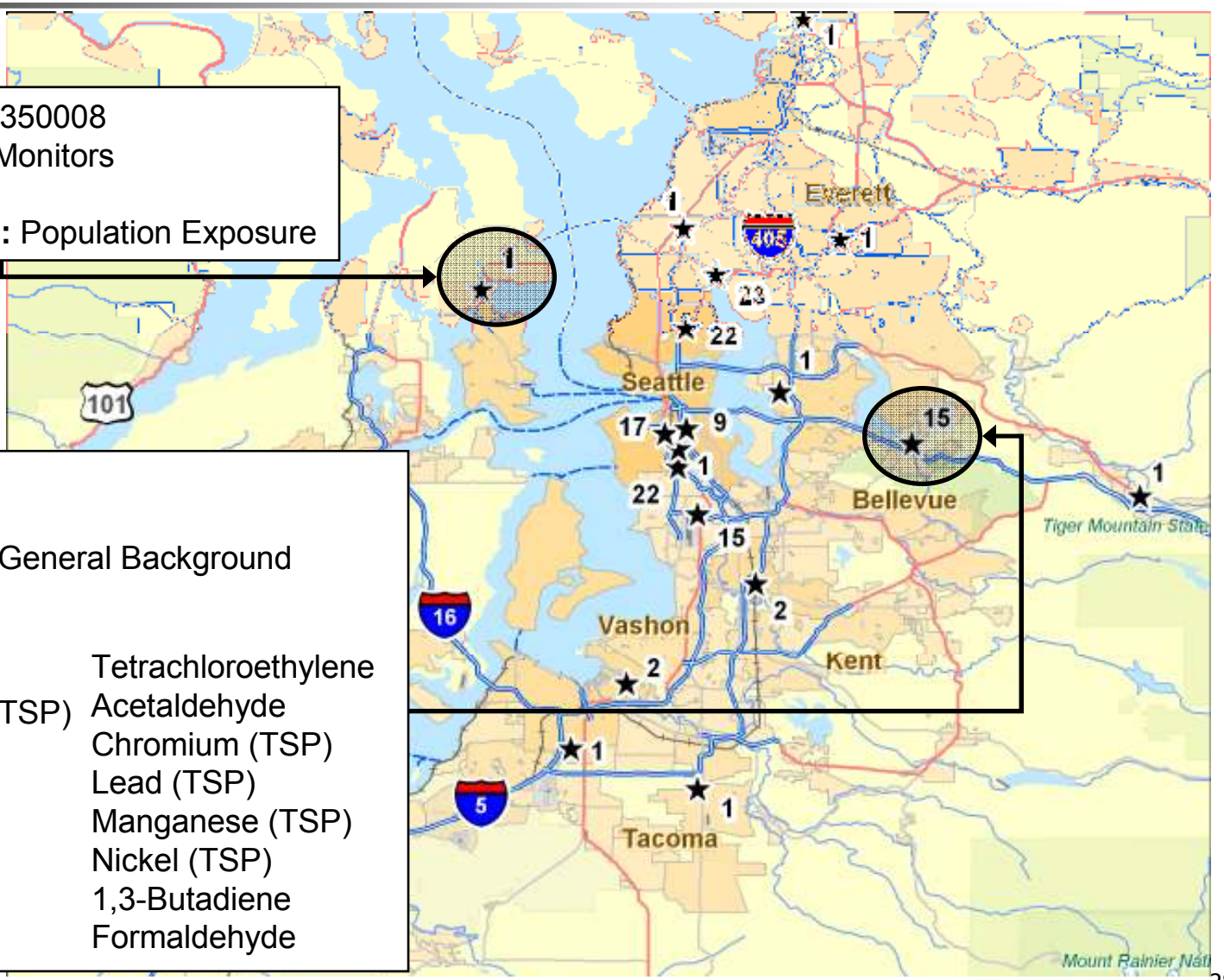
Parameters Monitored – Example

- Sum/tabulate the number of parameters measured at each monitor location
- Map created in ESRI ArcGIS Desktop, Version 9.1
- Data within the map includes
  - Geographic features from U.S. Census, Tele Atlas, National Park Service (all available from the ESRI data collection)
  - Air quality monitor locations and parameters from EPA's AQS

Greater Seattle area

Parameters Monitored – Example

$PM_{2.5}$ Monitor ID: 530350008
Monitor Type: Tribal Monitors
1999-Present
Monitoring Objective: Population Exposure



Monitor ID: 530330010
1975-Present
Monitoring Objective: General Background

Pollutants Measured:

Arsenic (TSP)
Suspended particulate (TSP)
Cadmium (TSP)
Chloroform
Carbon tetrachloride
Benzene
Trichloroethylene
Tetrachloroethylene
Acetaldehyde
Chromium (TSP)
Lead (TSP)
Manganese (TSP)
Nickel (TSP)
1,3-Butadiene
Formaldehyde

Parameters Monitored – Example

Criteria pollutants monitored

CO	NO2	O3	SO2	PM2.5	PM10	PB	Total Creteria	Site ID	City	County
1	1	0	0	0	0	0	0	530330032	Seattle	King Co
0	0	0	0	1	0	1	2	530330024	Lake Forest Park	King Co
0	0	0	0	0	0	1	1	530330038	Seattle	King Co
1	1	1	1	1	0	0	2	530330080	Seattle	King Co
0	0	1	0	0	0	1	1	530330010	Bellevue	King Co
0	0	0	0	0	0	1	1	530330020	Seattle	King Co
0	0	0	0	1	1	0	2	530330057	Seattle	King Co
0	0	0	0	1	1	0	2	530332004	Kent	King Co
0	0	0	0	1	1	0	2	530670013	Lacey	Thurston Co
0	0	0	0	1	1	0	2	530530031	Tacoma	Pierce Co
0	0	0	0	1	0	0	1	530330037	Bellevue	King Co
0	0	0	0	1	0	0	1	530611007	Marysville	Snohomish Co
0	0	0	0	1	0	0	1	530610005	Mountlake Terrace	Snohomish Co
0	0	0	0	1	0	0	1	530531018	Puyallup	Pierce Co
0	0	0	0	1	0	0	1	530330027	Redmond	King Co
0	0	0	0	1	0	0	1	530330021	Seattle	King Co
0	0	0	0	1	0	0	1	530450004	Shelton	Mason Co
0	0	0	0	1	0	0	1	530530029	Tacoma	Pierce Co
0	0	1	1	0	0	0	1	530090012		Clallam Co
0	0	1	0	1	0	0	1	530330017		King Co
0	0	0	0	1	0	0	1	530350008		Kitsap Co

Top 10 monitors based on the number of parameter measured

Site ID	City	County	Total Pollutants	RANK
530330024	Lake Forest Park	King	23	1
530330032	Seattle	King	22	2
530330038	Seattle	King	22	3
530330080	Seattle	King	17	4
530330010	Bellevue	King	15	5
530330020	Seattle	King	15	6
530330057	Seattle	King	9	7
530332004	Kent	King	2	8
530670013	Lacey	Thurston	2	9
530530031	Tacoma	Pierce	2	10

Rank of importance

HAP pollutants monitored

All HAPs	Site ID	Site Address	City	County	State
22	530330032	6431 Corson Ave S	Seattle	King Co	WA
21	530330024	Lake Forest Park Towne Center/Bothellway	Lake Forest Park	King Co	WA
21	530330038	8241 14th Ave. N.E.	Seattle	King Co	WA
15	530330080	Beacon Hill Reservoir/Charleston & 15th	Seattle	King Co	WA
14	530330010	Lake Sammamish State Park/20050 Se 56th	Bellevue	King Co	WA
14	530330020	2501 S 150th (Seatac North)	Seattle	King Co	WA
7	530330057	Duwamish Pump Sta/4752 E Marginal Wy S	Seattle	King Co	WA

Regional/Local Versus National Comparison

- Motivation:
  - Sites that measure high concentrations are important for assessing NAAQS compliance
  - Comparisons of nationwide data to monitors within a given network show whether certain sites are candidates for removal or repurposing
- Resources:
  - Concentration data from AQS or EPA AIRNow Tech
  - Statistical software and GIS may be helpful, but are not necessary

Regional/Local Versus National – Example

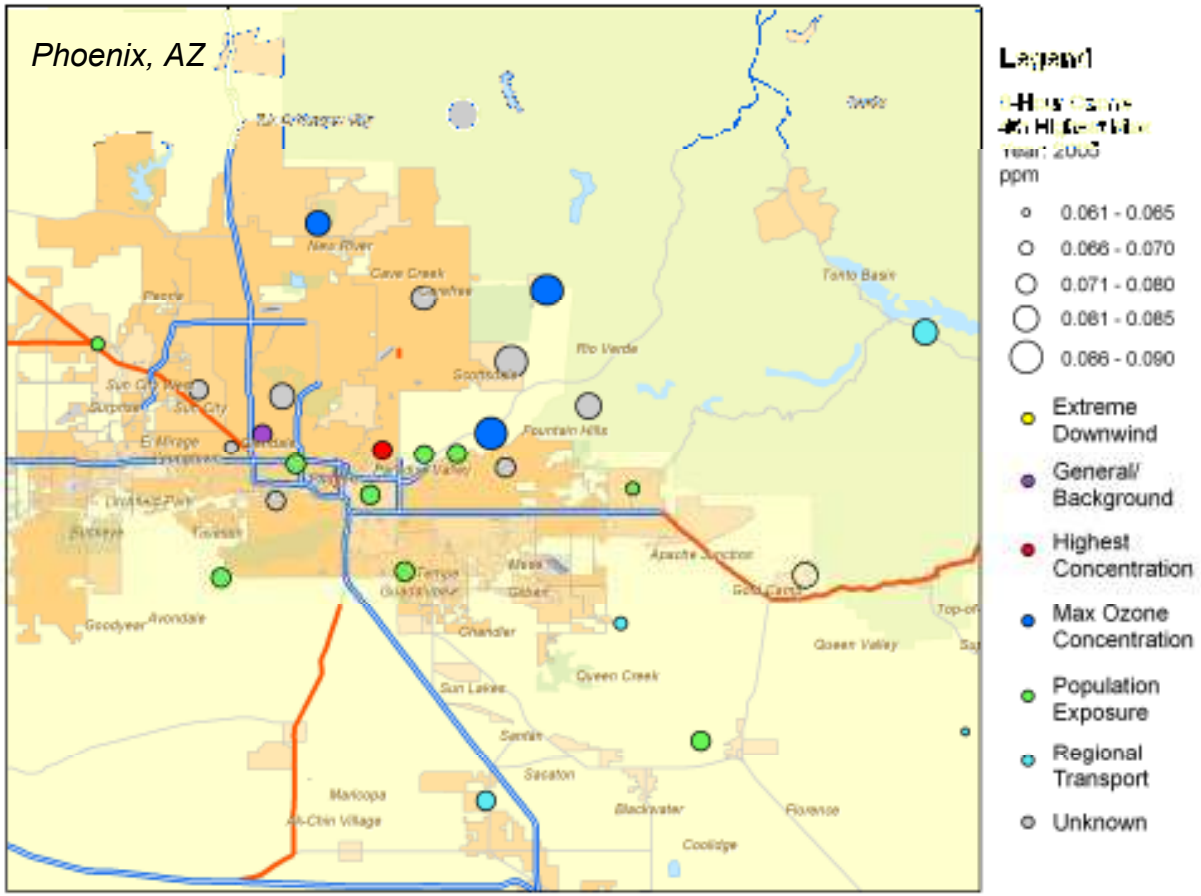
- Compare monthly average 8-hr ozone and the national monthly average 8-hr ozone site by site in an MSA, region, etc.
- Which, if any, monitoring sites within the selected domain correlate/differ from the national average?
  - Do the monitoring objectives support this comparison?
  - Where are the monitors located? Are the geographic surroundings unique?

Regional/Local Versus National – Example

Gather site level monitor values from AQS in order to make comparisons with national data – example is for Phoenix, Arizona

	PM2.5 (µg/m3)												
	24-Hour Values							Annual					
Row #	# Obs	1st Max	2nd Max	3rd Max	4th Max	98th Pct	# Exceed	Mean	# Exceed	Year	Site ID	Site Address
SORT												
1	87	57	37	35	29	37	0	14.8	0	2004	010735002	Pinson, High Sch., Box 360 Hwy 75 North
2	272	53	50	47	47	44	0	18.0	1	2004	010730023	No. B'Ham,Sou R.R., 3009 28th St. No.
3	89	53	34	34	30	34	0	13.8	0	2003	010690002	East Highland St., Board Of Ed. Bldg.
4	358	53	43	42	39	35	0	15.6	1	2003	010732003	1242 Jersey St Wylam Al
5	40	50	42	33	31	50	0	17.9	1	2004	010732003	1242 Jersey St Wylam Al

EPA AirData: Monitor Values Report - Criteria Air Pollutants



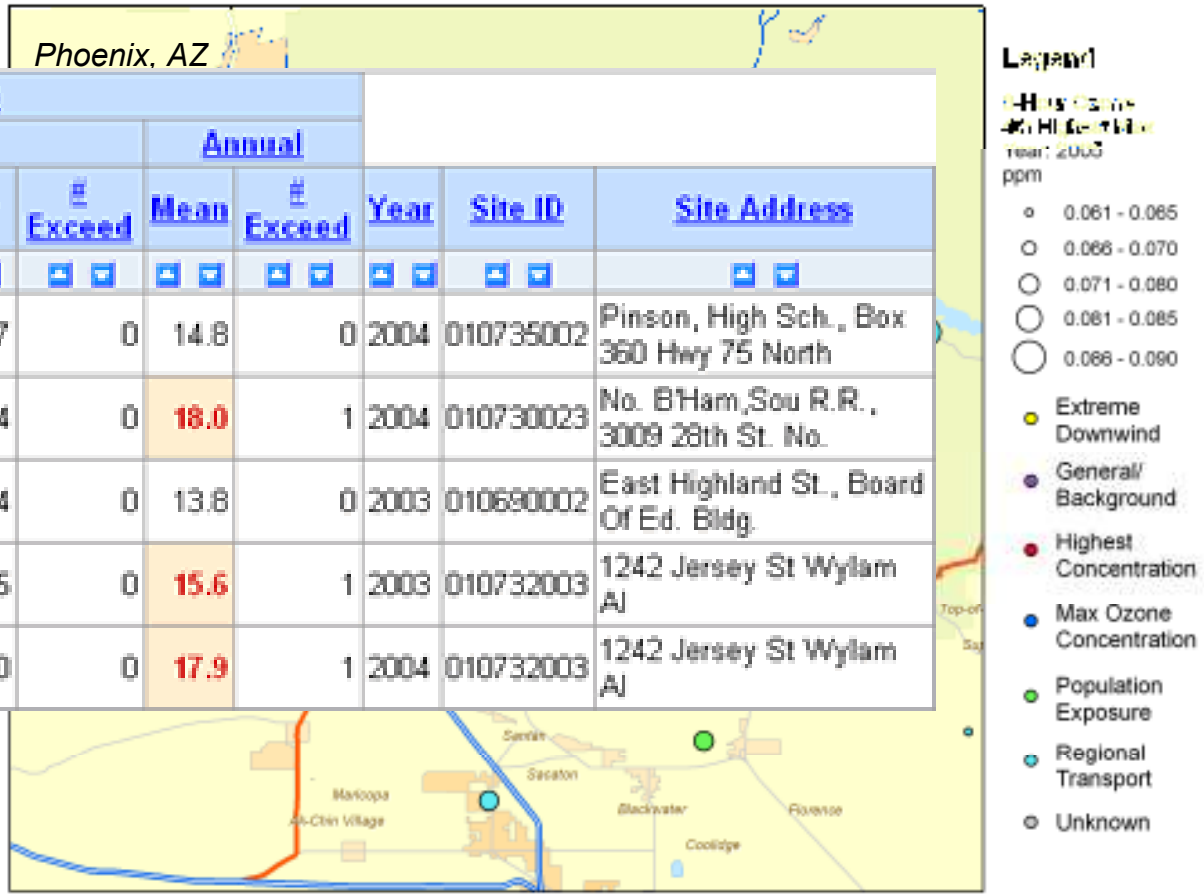
Graphic display of monitor values and monitor objectives

Regional/Local Versus National – Example

Gather site-level monitor

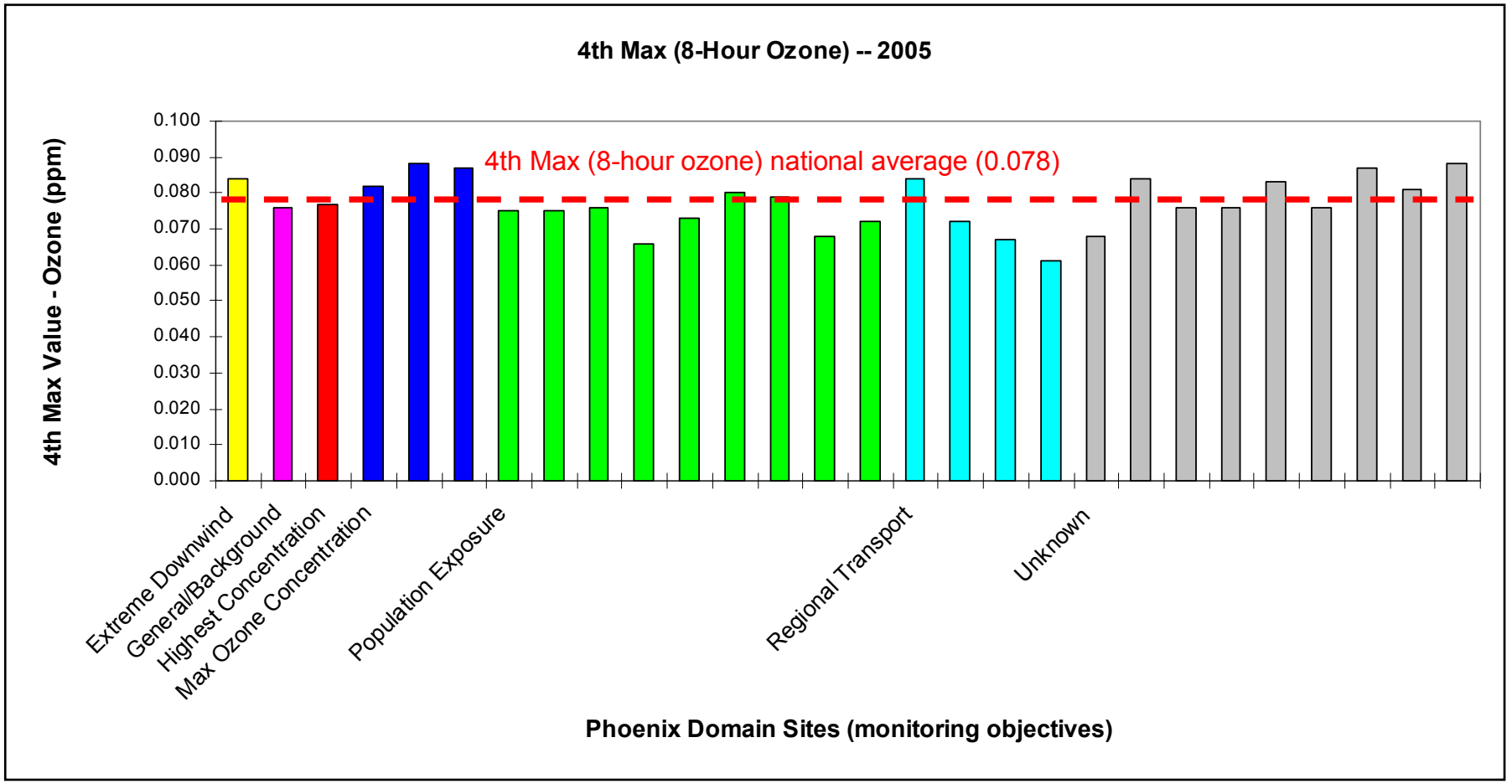
	PM2.5 (µg/m3)											
	24-Hour Values							Annual				
Row #	# Obs	1st Max	2nd Max	3rd Max	4th Max	98th Pct	# Exceed	Mean	# Exceed	Year	Site ID	Site Address
SORT												
1	87	57	37	35	29	37	0	14.8	0	2004	010735002	Pinson, High Sch., Box 360 Hwy 75 North
2	272	53	50	47	47	44	0	18.0	1	2004	010730023	No. B'Ham,Sou R.R., 3009 28th St. No.
3	89	53	34	34	30	34	0	13.6	0	2003	010690002	East Highland St., Board Of Ed. Bldg.
4	358	53	43	42	39	35	0	15.6	1	2003	010732003	1242 Jersey St Wylam Al
5	40	50	42	33	31	50	0	17.9	1	2004	010732003	1242 Jersey St Wylam Al

EPA AirData: Monitor Values Report - Criteria Air Pollutants

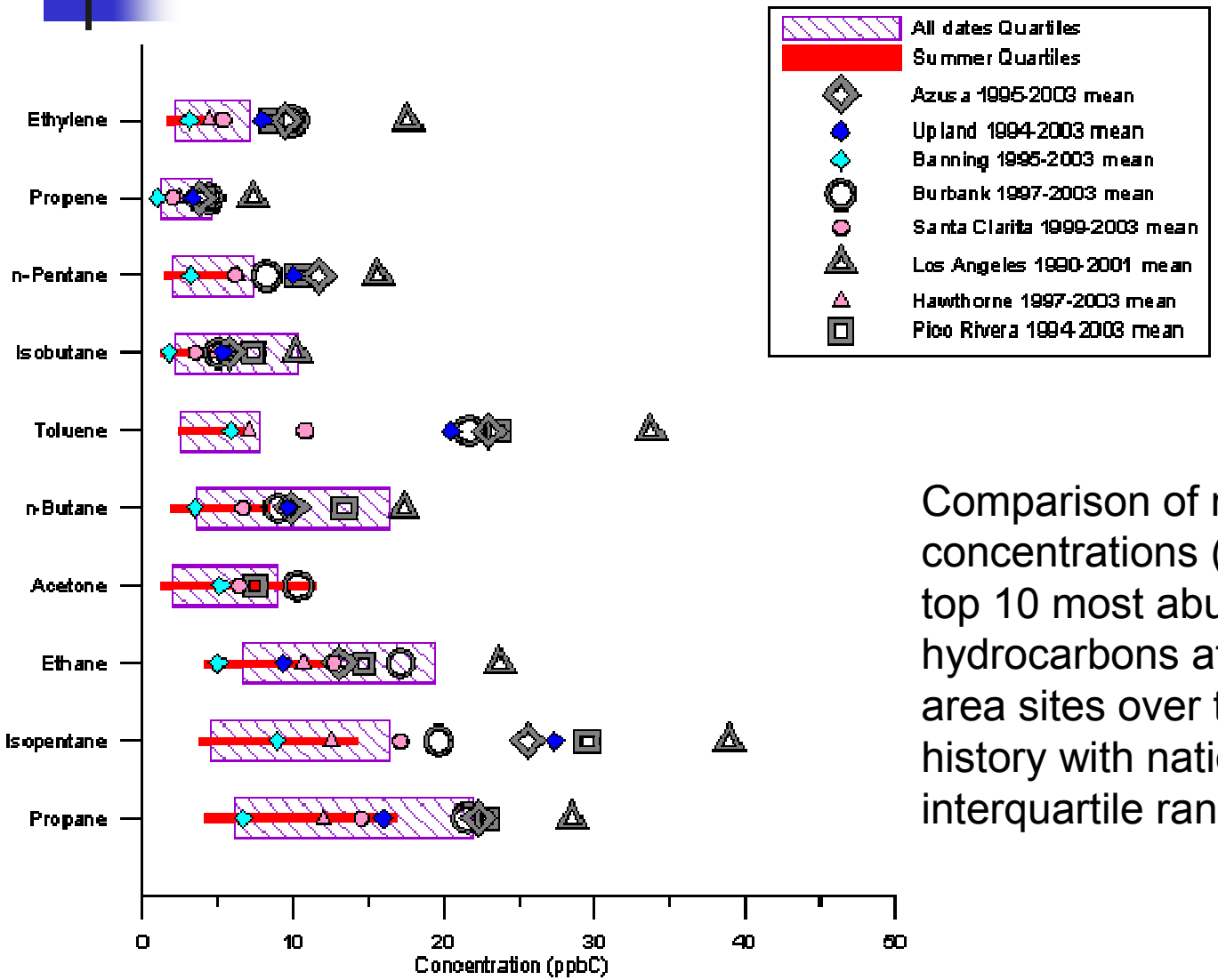


Graphical display of monitor values and monitor objectives

Regional/Local Versus National – Example



Regional/Local Versus National – Example



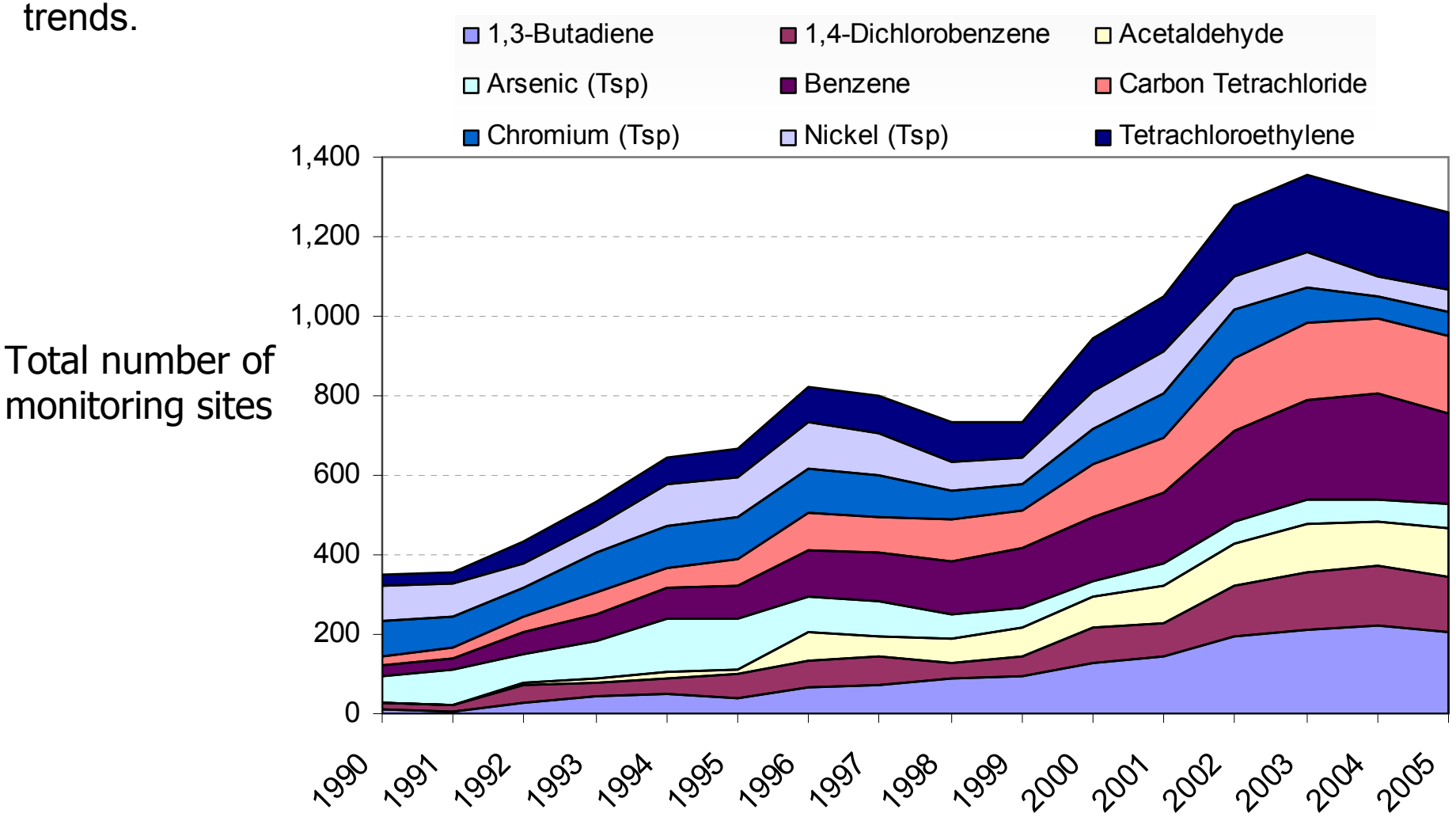
Comparison of mean concentrations (ppbC) of the top 10 most abundant hydrocarbons at Los Angeles area sites over their operating history with national interquartile ranges.

Trends Impacts

- Motivation:
  - Monitors that have long historical trends are valuable for tracking trends
  - This technique places the most importance on sites with the longest continuous trend record
- Resources needed:
  - Historical monitor data from AQS or EPA AIRNow Tech
  - Concentration data may be helpful, but are not necessary

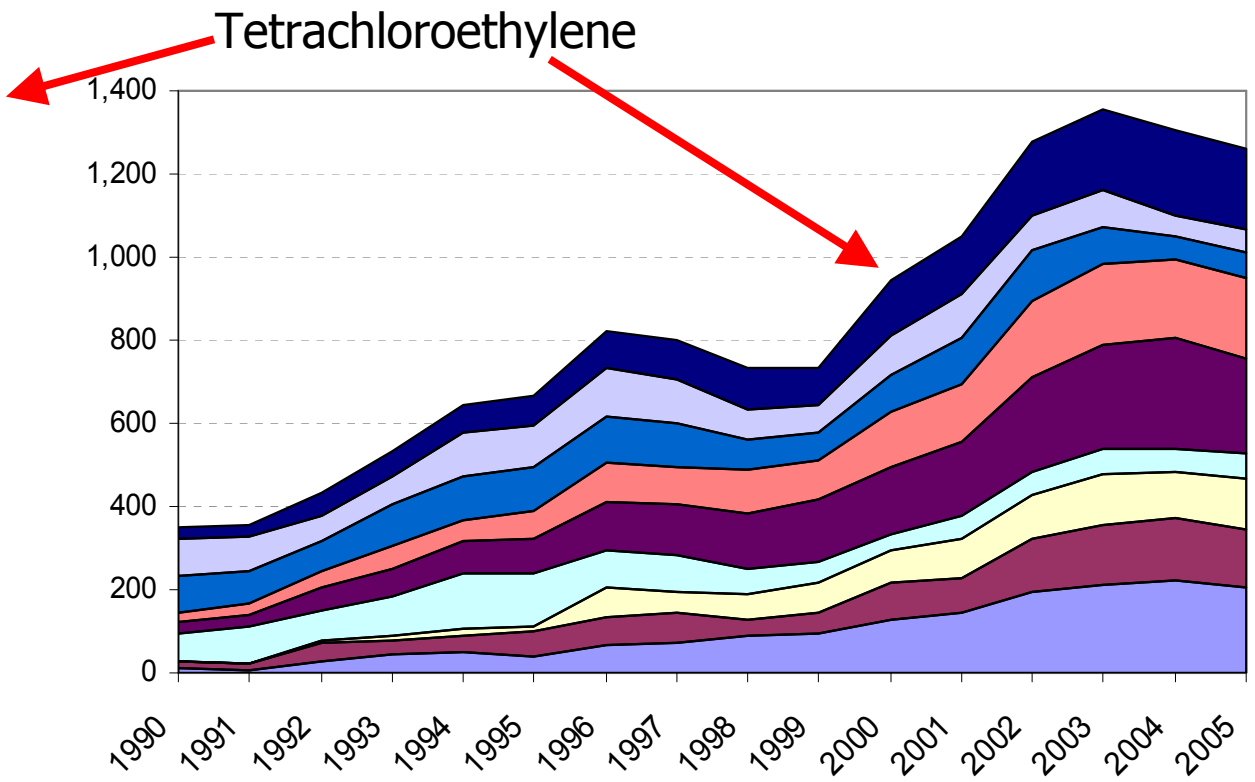
Trends Impacts – Examples

Monitors in the United States that have long historical trends are valuable for tracking trends.



Trends Impacts – Examples

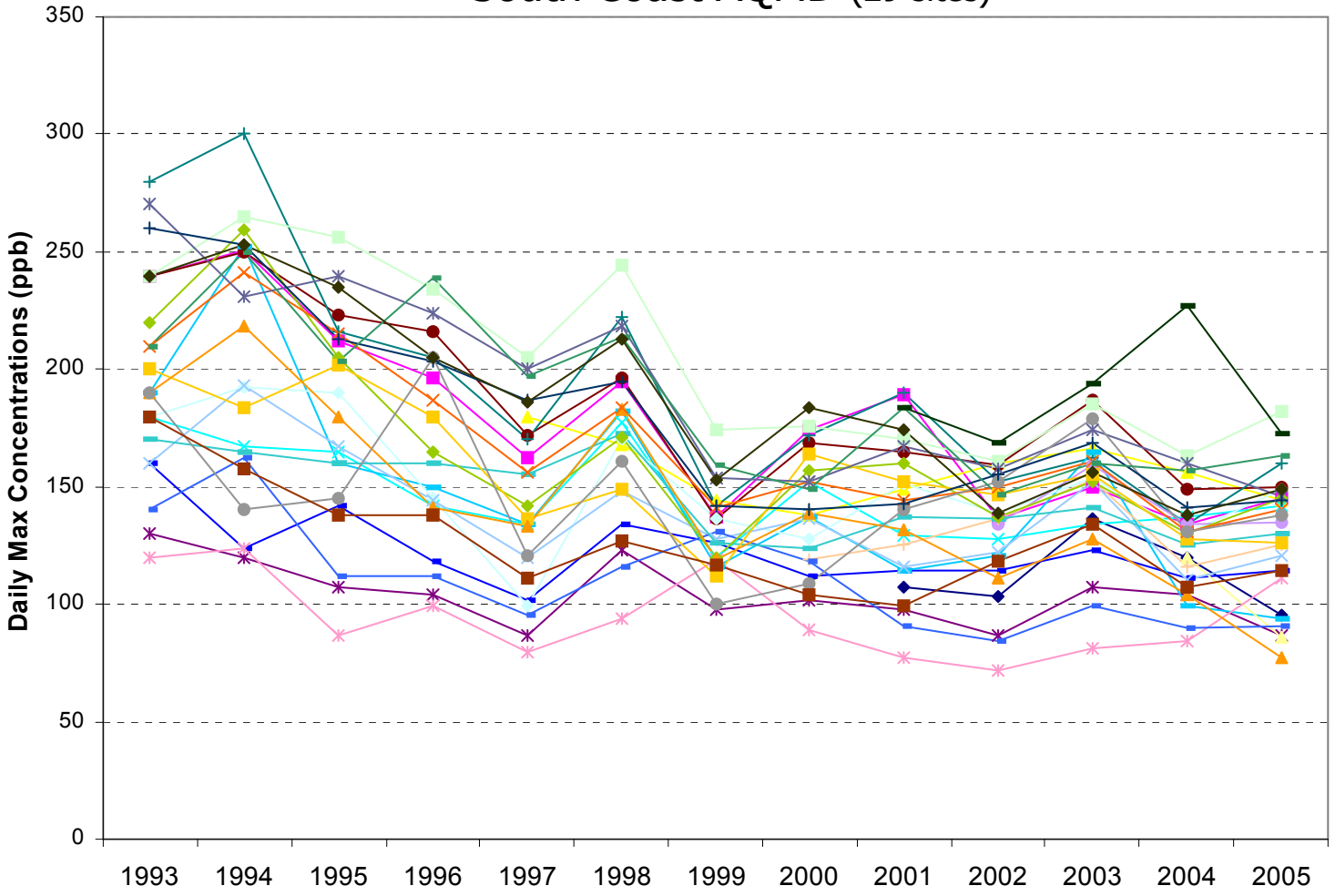
City, State	AQS SiteID	Years
Stockton, CA	06-077-1002	13
Baltimore, MD	24-510-0040	12
Los Angeles, CA	06-037-1002	11
San Francisco, CA	06-001-1001	10
Fresno, CA	06-019-0008	10
Baltimore, MD	24-005-3001	10
Los Angeles, CA	06-037-1103	9
Los Angeles, CA	06-037-4002	9
San Diego, CA	06-073-0003	9
San Francisco, 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



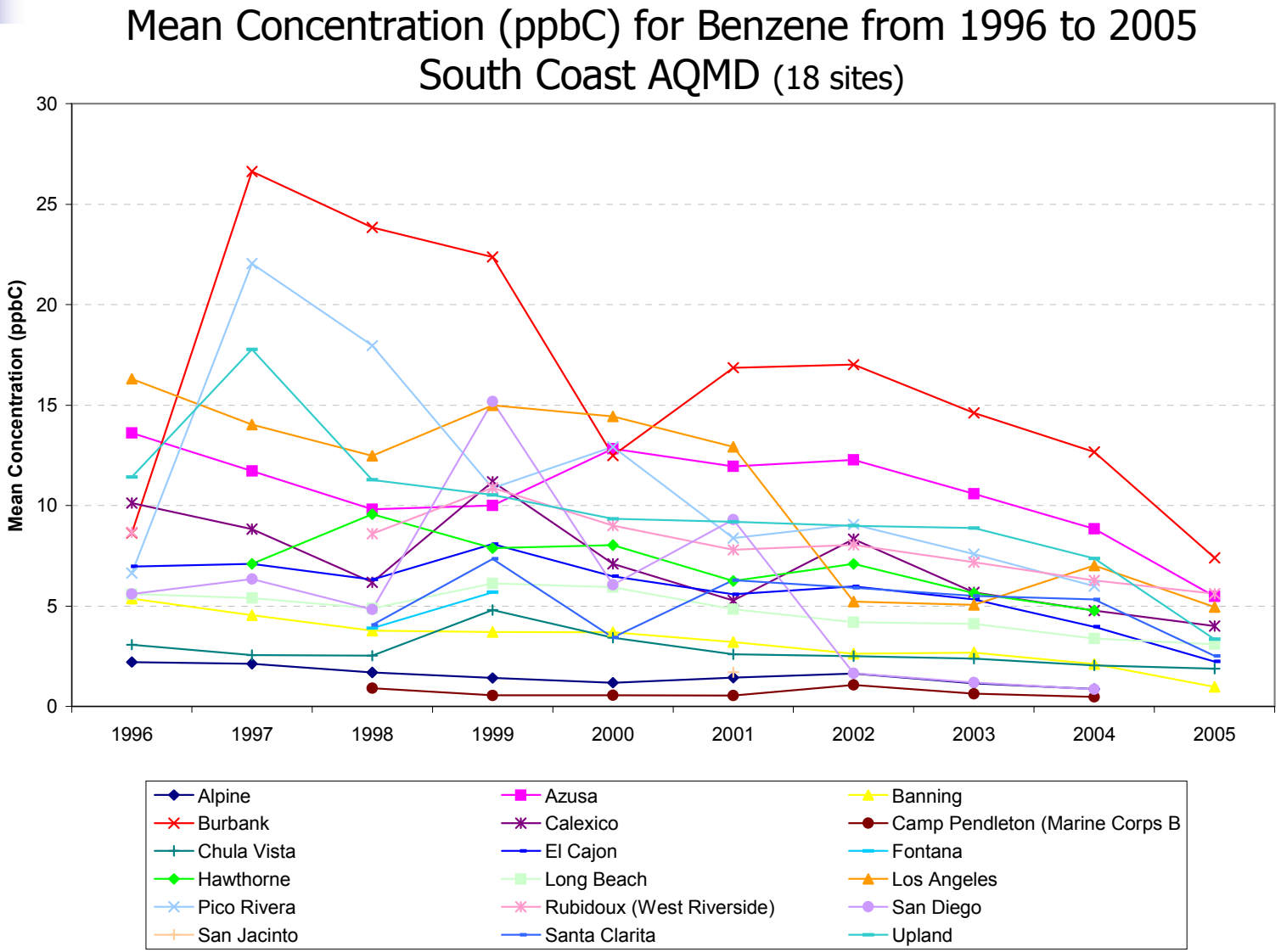
Number of annual averages available for tetrachloroethylene at toxics trends sites from 1990 to 2003. For this analysis, sites with the longest record would be rated higher than those with shorter records.

Trends Impacts – Examples

Daily Maximum Concentration (ppb) for Ozone 1-hr from 1993 to 2005
South Coast AQMD (29 sites)



Trends Impacts – Examples



Measured Concentration

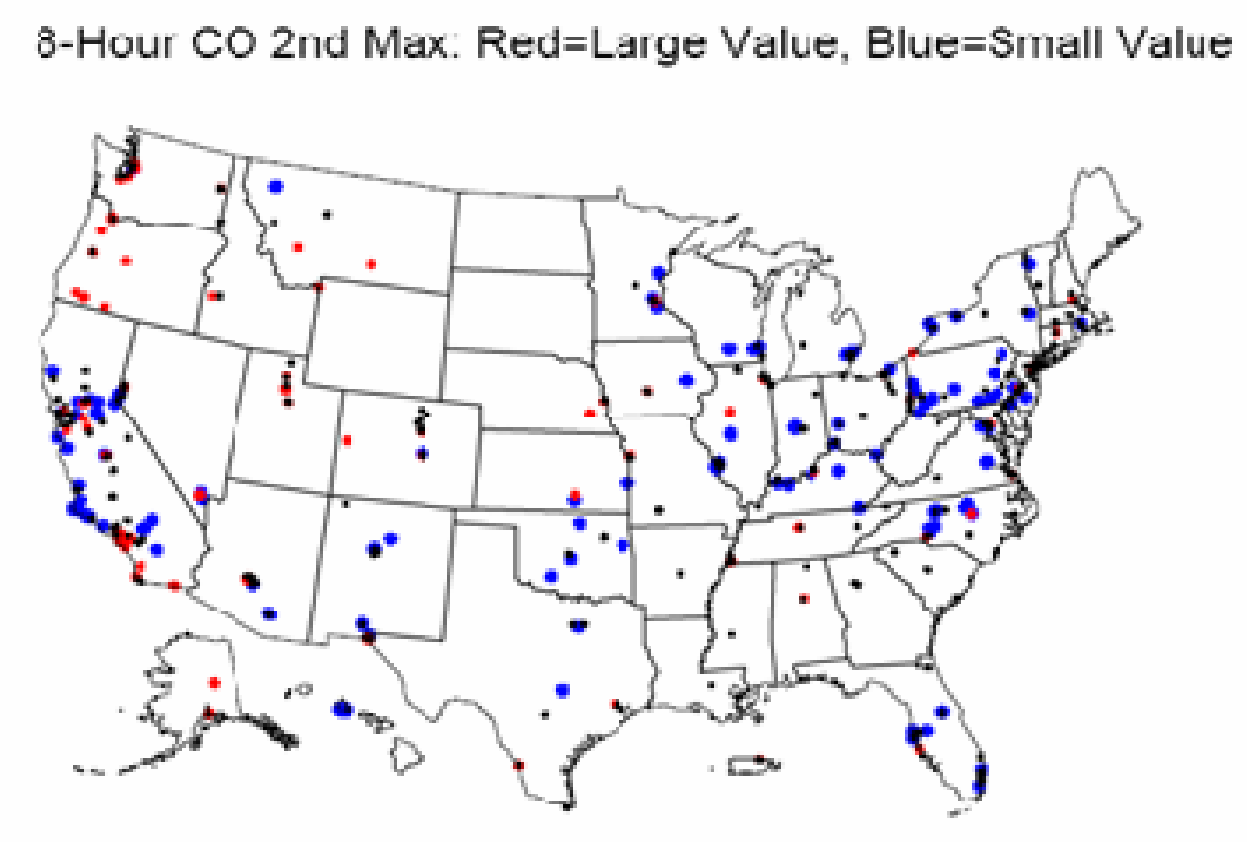
- Motivation:
  - Individual sites are ranked based on the concentrations of pollutants they measure
  - Results can be used to determine which monitors are less useful in meeting the selected objective
- Resources needed:
  - Concentration data from AQS or EPA AIRNow Tech
  - Statistical software, detailed site information, and GIS may be helpful, but are not necessary

Measured Concentration – Goals

- Sites that measure high concentrations are important for assessing NAAQS compliance and population exposure (AQI) 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.

Measured Concentration – Example

- This metric was one of five used in the 2000 National Analysis.
- Sites in red record the highest CO concentrations and are the most valuable.
- Sites in blue record the lowest values and are candidates for removal or repurposing.



Schmidt M. (2001) Monitoring strategy: national analysis

Deviation from NAAQS

- Motivation:
  - 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.
- Resources needed:
  - Concentration data from AQS or EPA AIRNow Tech
  - Site locations, historical data, and GIS may be helpful, but are not necessary

Deviation from NAAQS – Goals (1 of 2)

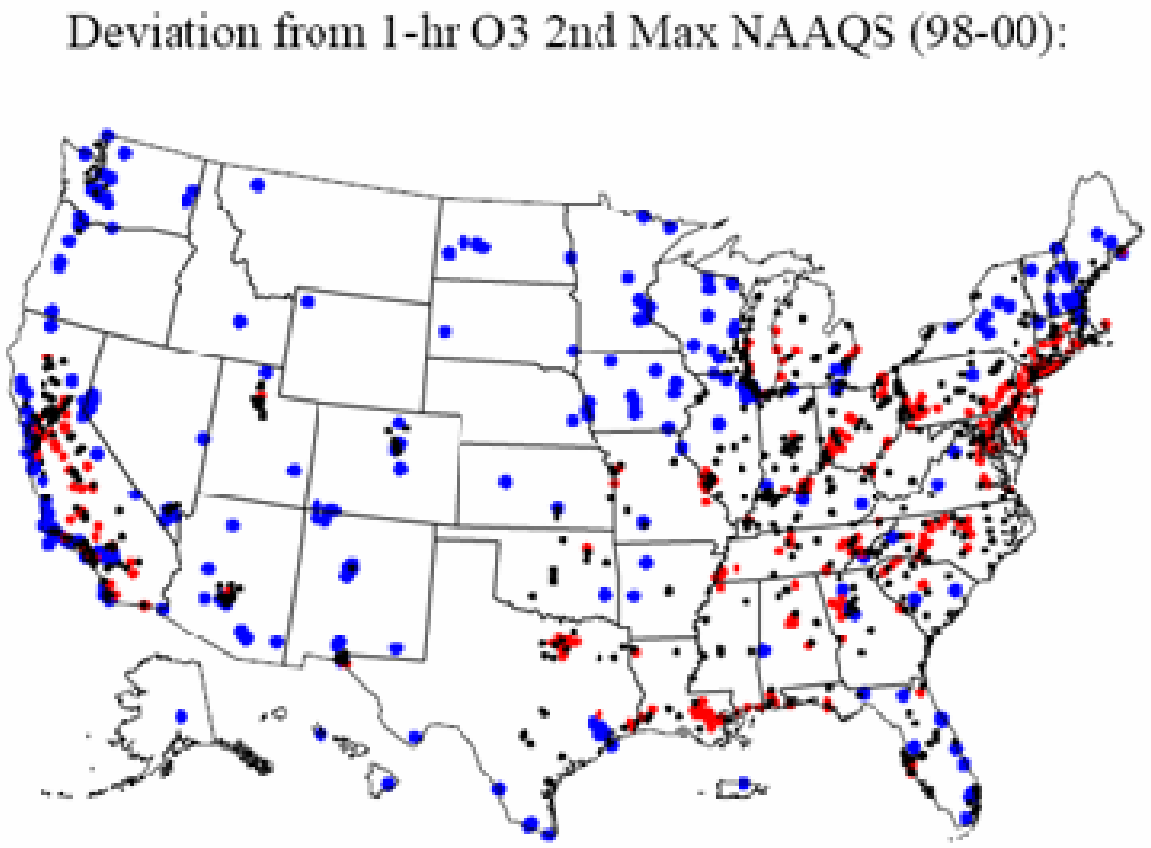
- This technique contrasts the difference between the standard and actual measurements or design values.
- If a pollutant (e.g., annual and 24-hr average) has more than one standard, monitors can be scored for each standard.
- The absolute value of the difference between the measured design value and the standard can be used to score each monitor.

Deviation from NAAQS – Goals (2 of 2)

- 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.

Deviation from NAAQS – Example

- This metric was one of five used in the 2000 National Analysis.
- Red circles denote sites that are closest to the standard. These sites are ranked highest in this analysis.
- Blue circles are those well above or below the standard. These sites are candidates for removal or repurposing.
- Black circles are not well above, below or close to the standard.



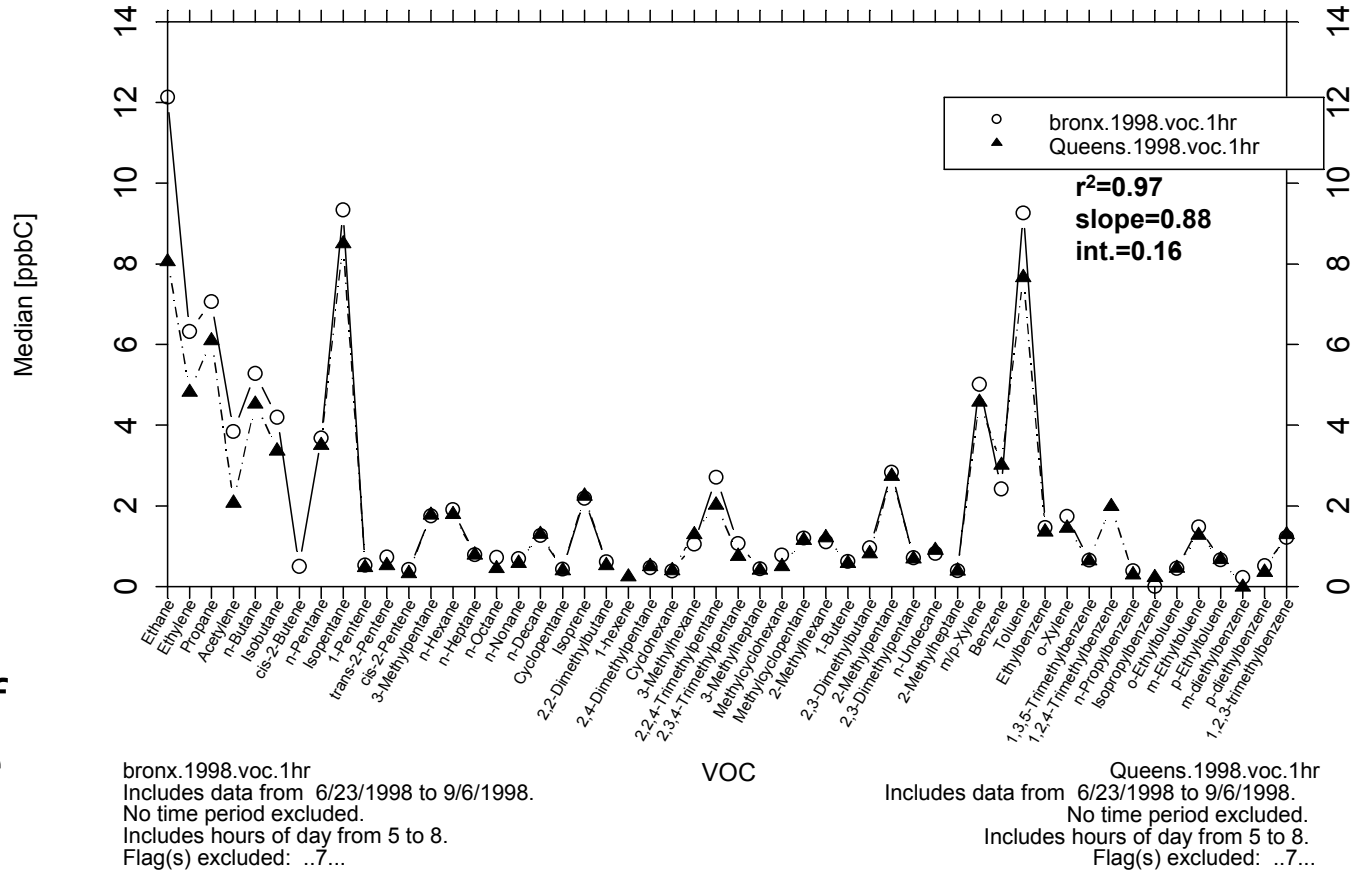
Schmidt M. (2001) Monitoring strategy: national analysis

Monitor-to-Monitor Correlation

- Motivation:
  - 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
- Resources needed:
  - Concentration data from AQS or EPA AIRNow Tech
  - Site locations, historical data, and GIS may be helpful, but are not necessary

Monitor-to-Monitor Correlation – Example

- Speciated hydrocarbon data were compared from two PAMS sites.
- Concentrations and composition compared well indicating one of the sites may be redundant.



Note that high correlation may exist in ranges of concentrations; it is important to evaluate correlation above certain levels, as these days may be driving NAAQS decisions.

Emission Inventory

- Motivation:
  - 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
- Resources needed:
  - County-level emission inventory data from the EPA National Emission Inventory (NEI) database (easily accessible from the EPA AirData web page)
  - County FIPS codes and/or geographic locations of monitor sites
  - A GIS to make simple county-level emission maps

FIPS = Federal Information Processing Standards Codes

Emission Inventory – Data Sources

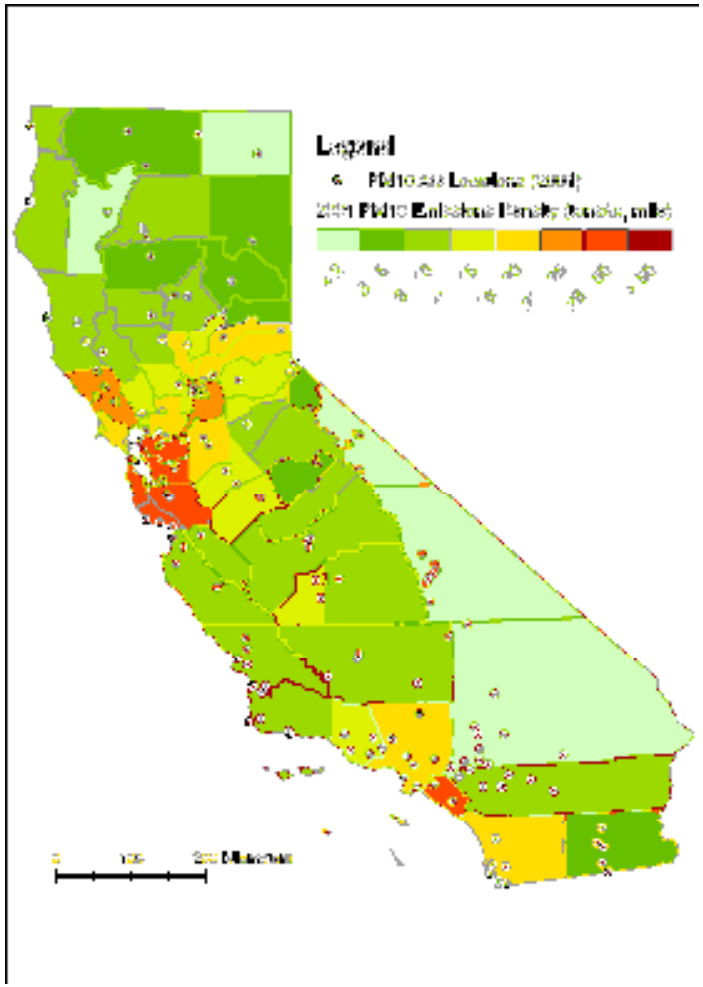
Emissions data available from EPA AirData web page:

- County-level CO, NO_x, VOC, SO_2, $PM_{2.5}$, PM_{10} or NH_3
- SIC based facility emissions for the pollutants listed above
- County-level Hazardous Air Pollutants (HAPs)
- SIC-based facility HAP emissions

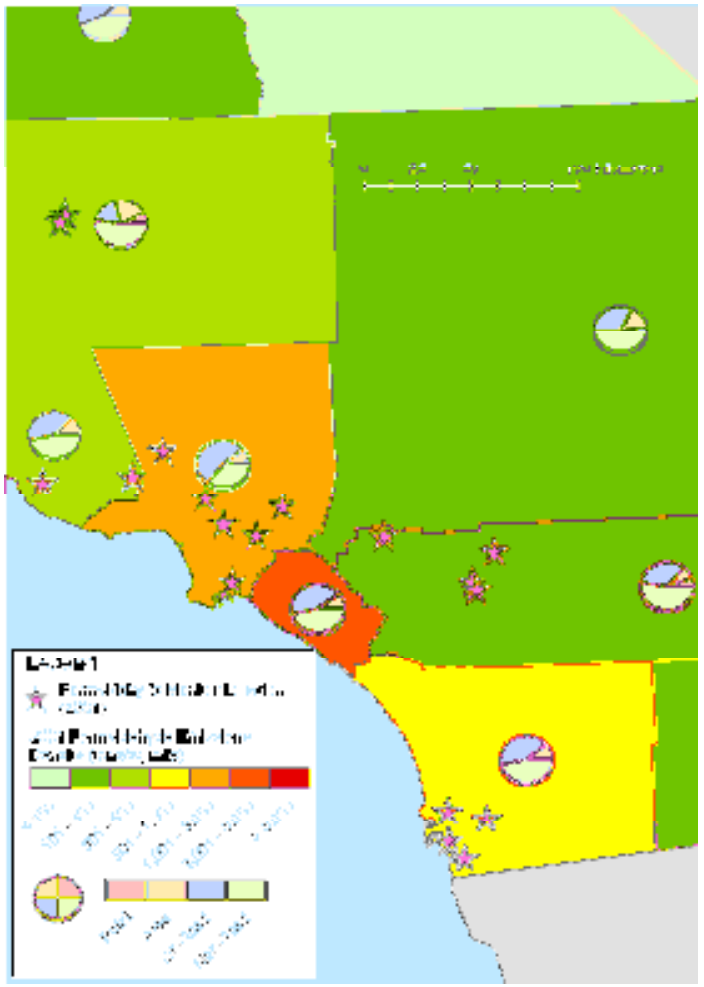
SIC = Standard Industrial Classification Code

Emission Inventory – Example

How is monitor coverage in comparison with PM_{10} emissions?

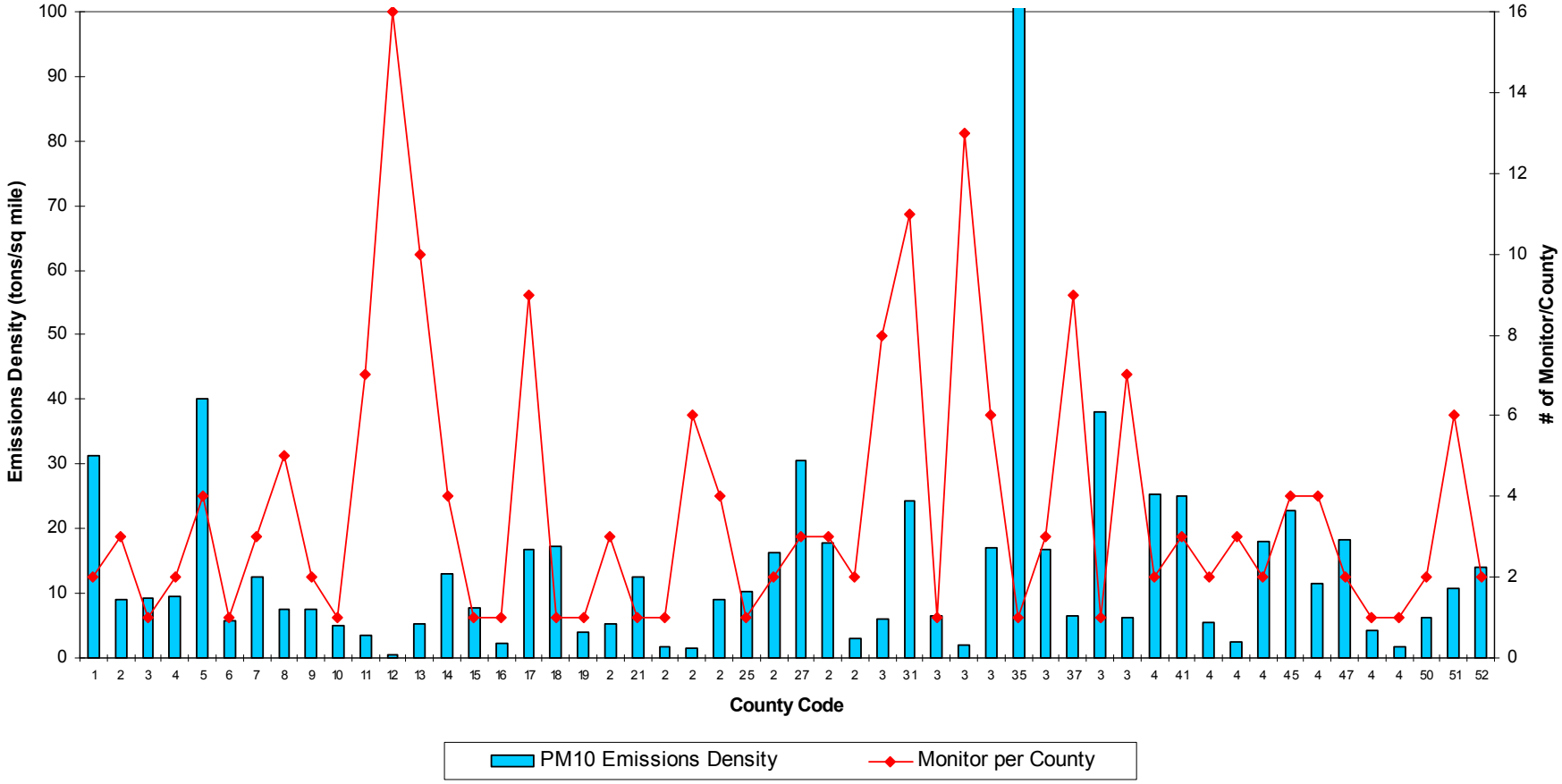


How does the coverage fare when looking at Formaldehyde emissions?



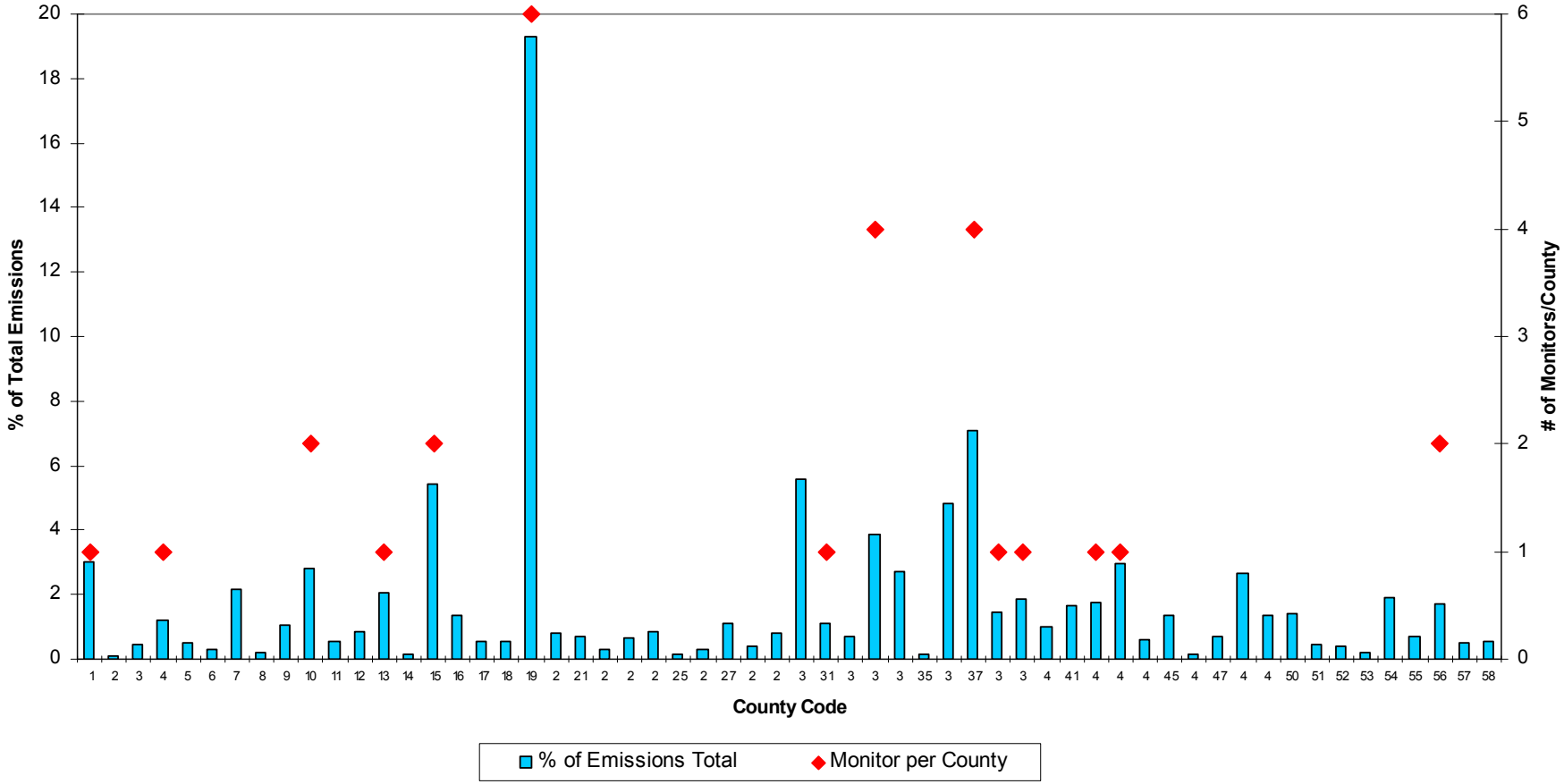
Emission Inventory – Example

Analysis technique allows further analysis on counties with high emissions and limited amount of monitors and vice versa.



Emission Inventory – Example

Analysis technique also depicts individual HAPs and where potential monitoring can be further investigated, based on emissions.



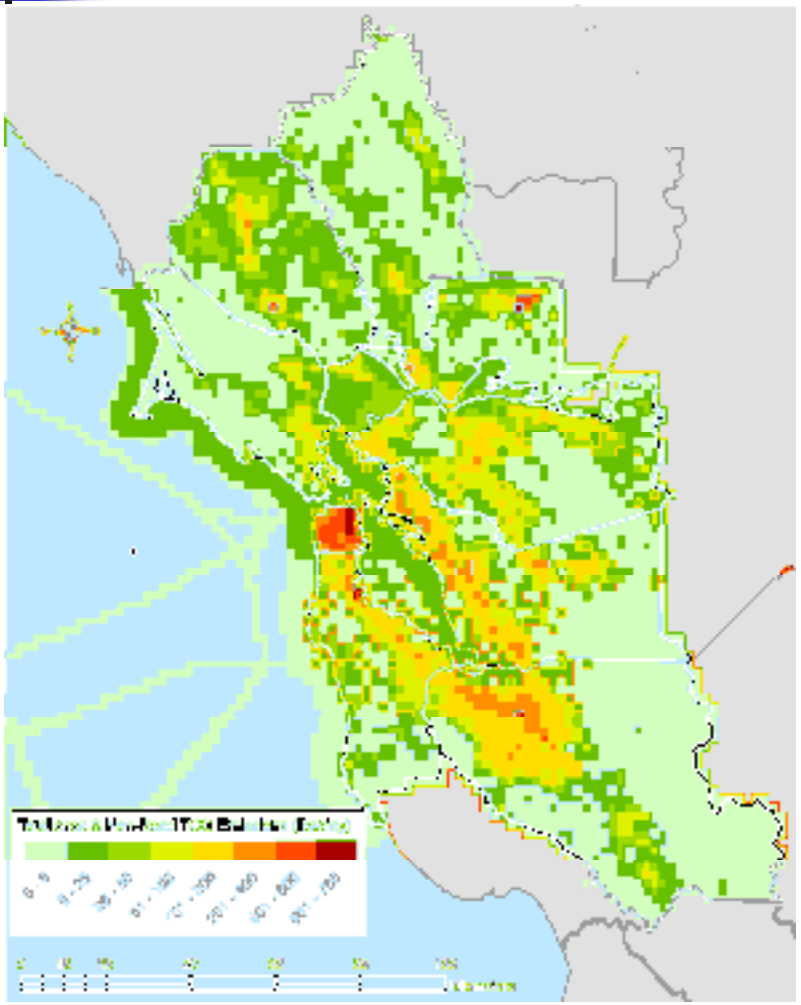
Gridded/Speciated Emission Inventory

- Motivation: Find locations where emissions of toxic pollutants are concentrated
- Can be used for any size network
- Various levels of complexity, depending on resources
  - The simplest version looks at county-level emissions of a single pollutant
  - More complex methods use gridded and/or species-weighted emissions
- Requires an emission inventory and GIS (if developing a gridded inventory)

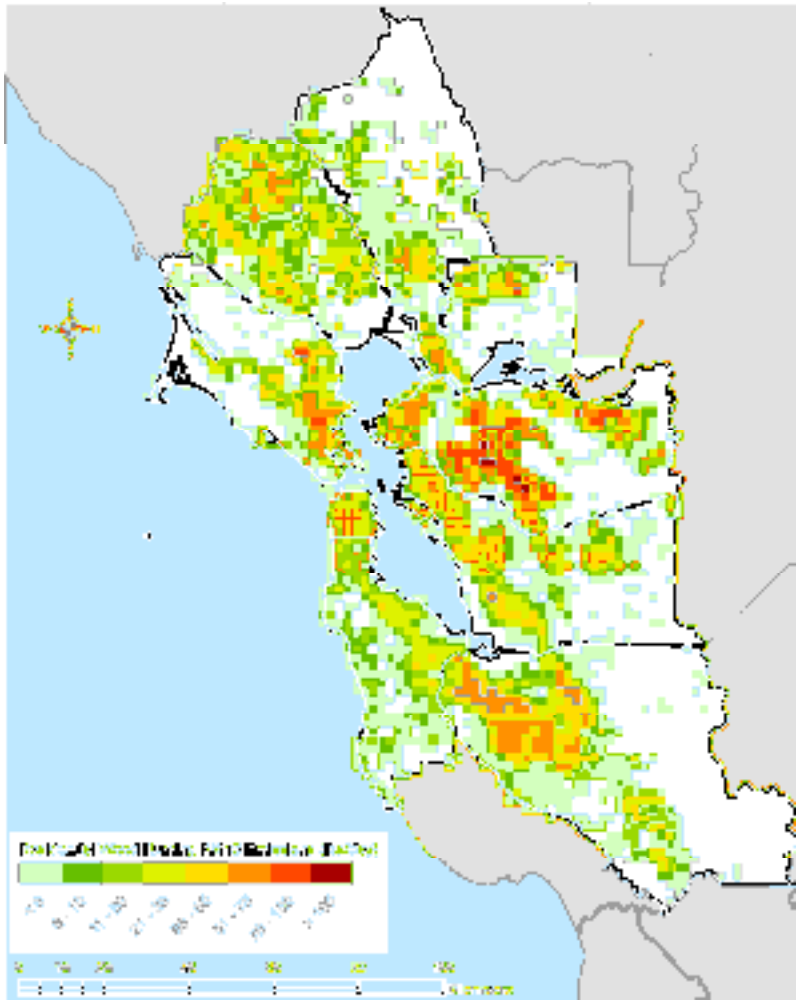
Training Example: Preparation of Gridded Emission Inventories of Toxic Air Contaminants for the San Francisco Bay Area (2006)

Funded by the Bay Area Air Quality Management District

BAAQMD Gridded Inventory Development

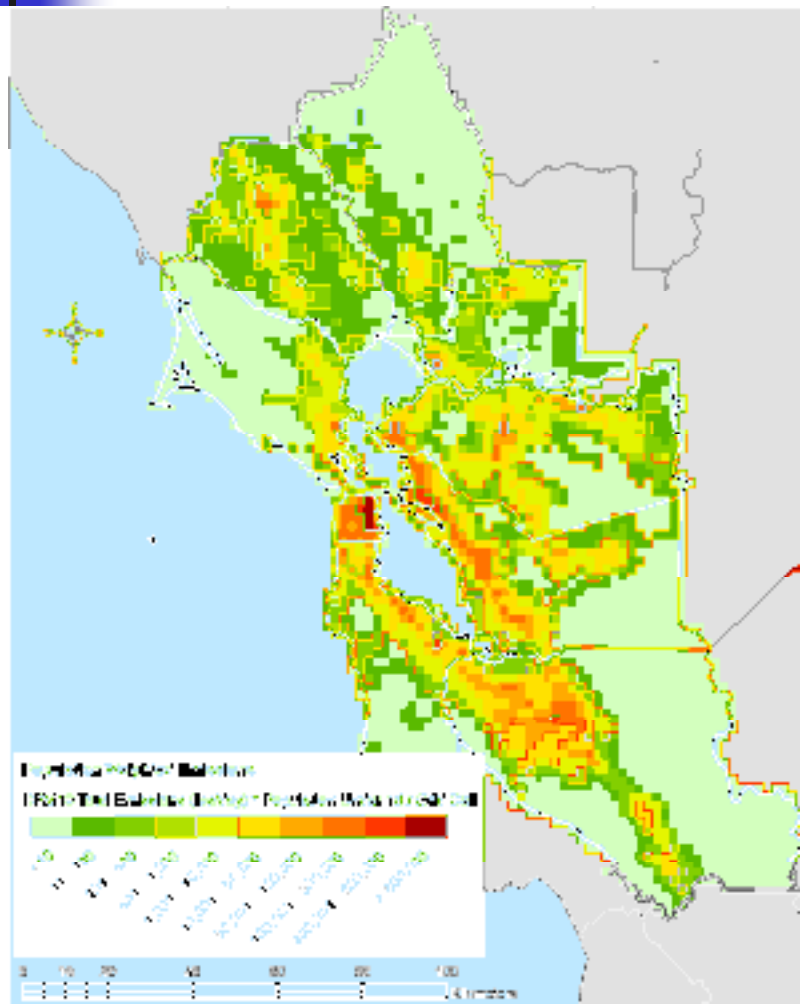


Total Area & Non-Road
TOG Emissions

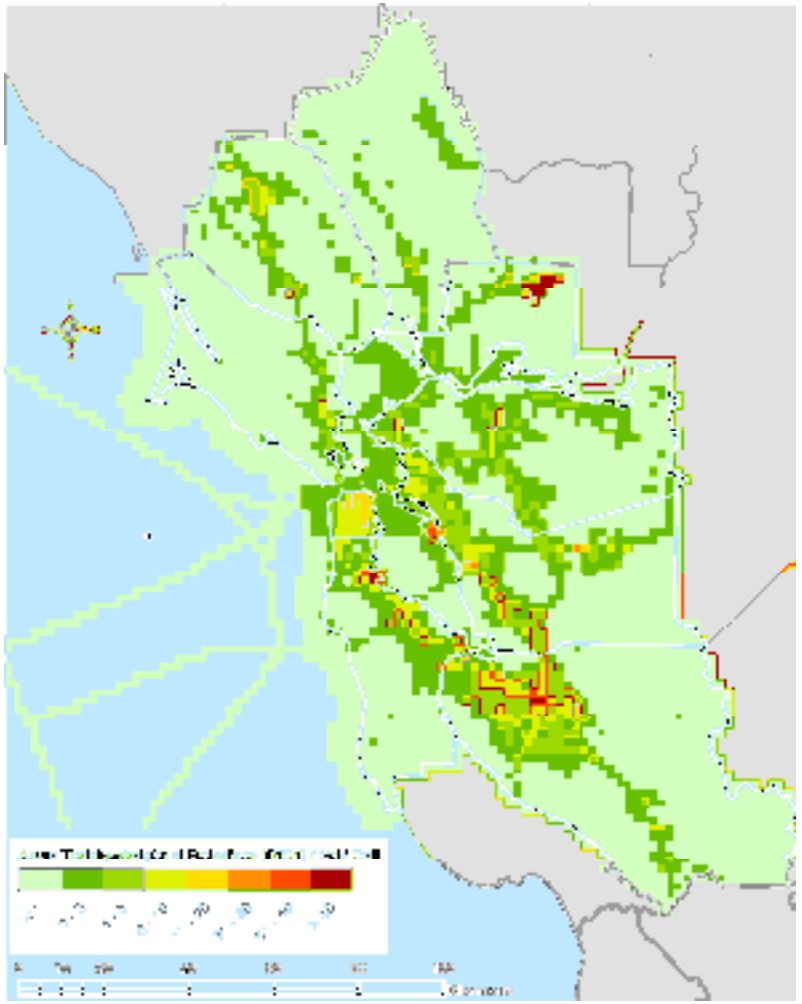


Residential Wood Burning
PM_{10} Emissions

BAAQMD Gridded Inventory Results

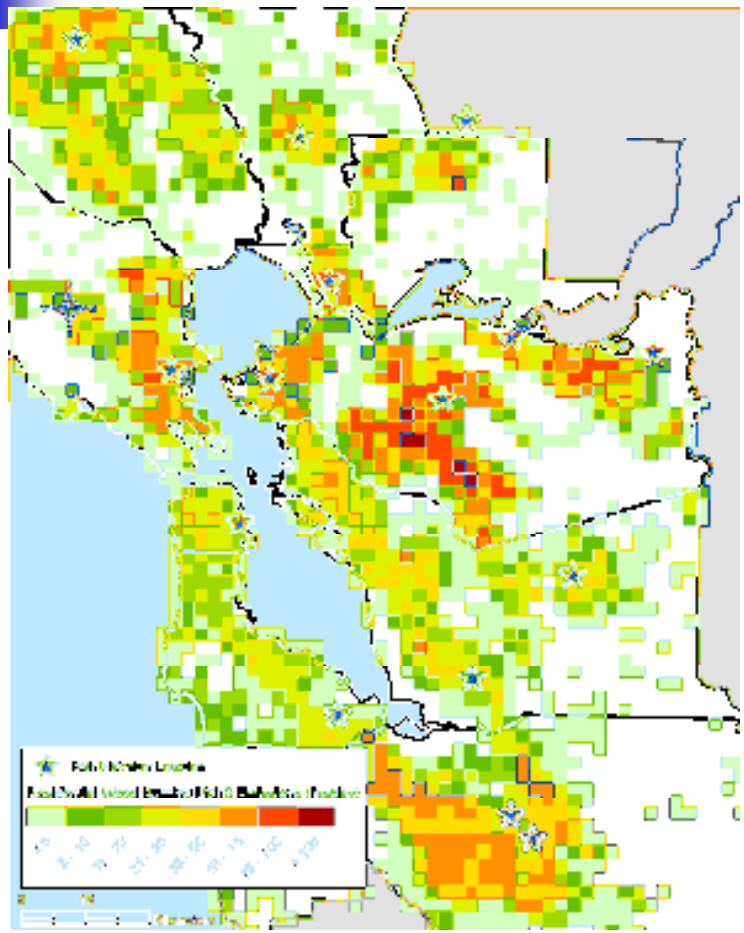


Total DPM emissions weighted by population under the age of 18

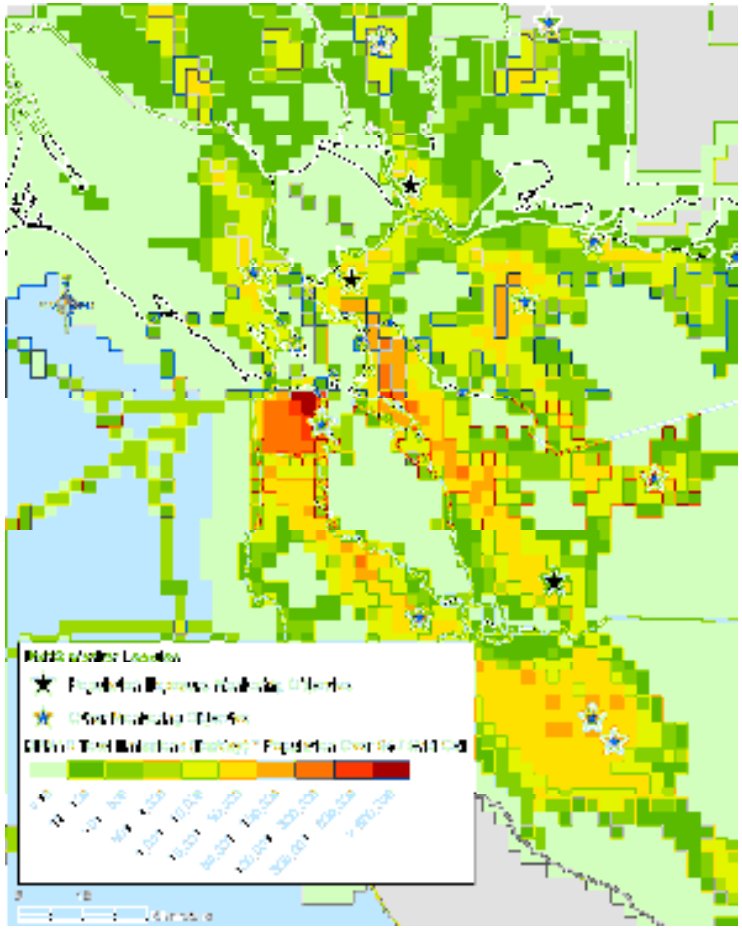


Acute toxicity-weighted emissions by grid cell

BAAQMD Gridded Inventory Results



PM_{10} monitors overlaid PM_{10} residential wood burning gridded emissions



Gridded DPM emissions weighted by sensitive populations. How do the population exposure monitors compare? Do other areas warrant additional monitors?

Gridded Emission Inventory – Conclusions

Gridded/speciated emission inventories can be used to

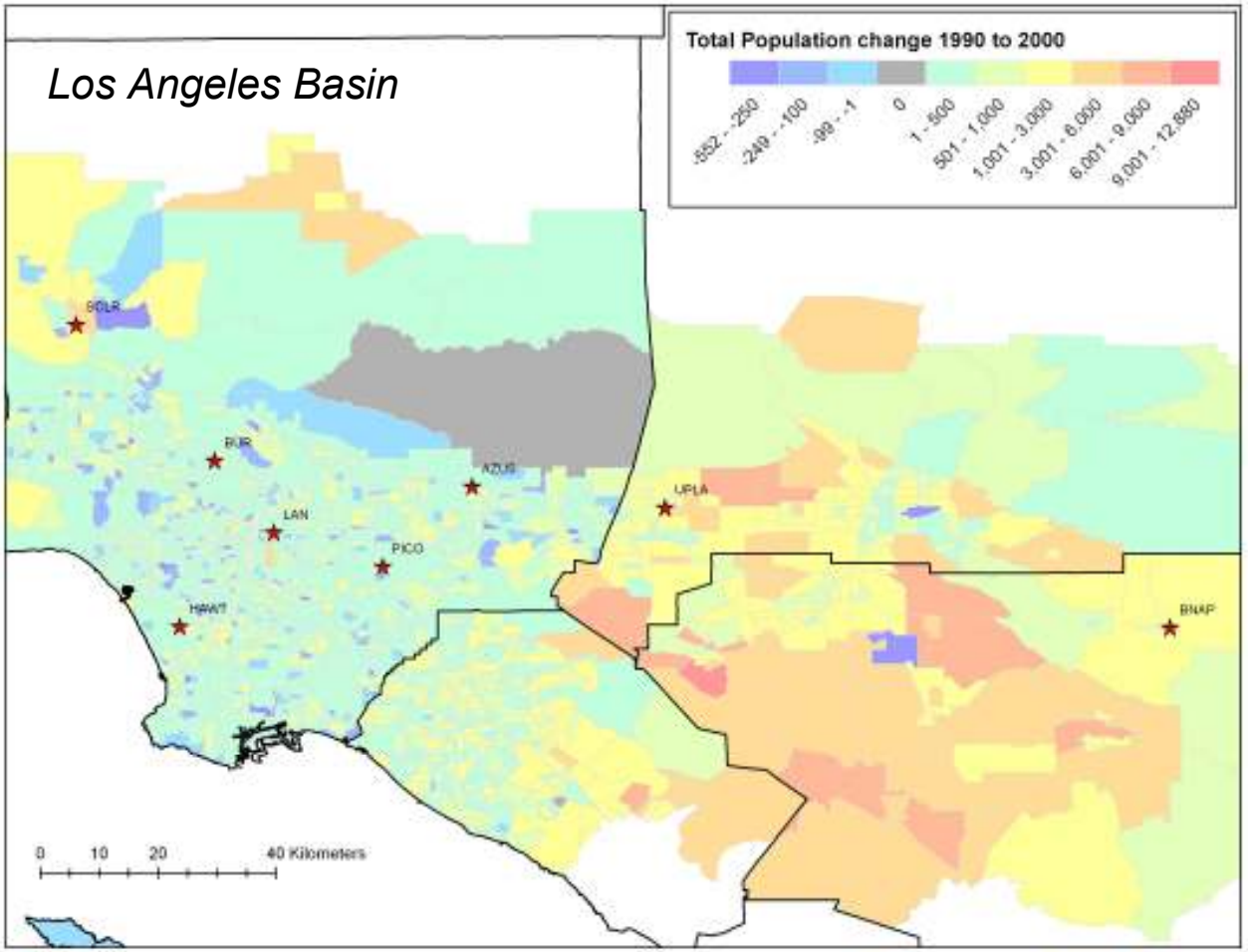
- Find areas at the grid cell level where emission concentrations are likely to be high
- Overlay existing monitor locations and see how they compare to areas of high emissions
- Select locations for new monitors
- Set priorities for monitoring
- Investigate a range of monitoring objectives and considerations

Addressing Population

- Motivation:
  - Need to understand if monitors are in areas of high population or if high rates of population change are associated with increased potential emissions activity and exposure
- Resources needed:
  - Sub-county level population data (current and historical) from the U.S. Census Bureau
  - Geographic monitor locations
  - Geographic Information System (GIS)

Population Change – Approach

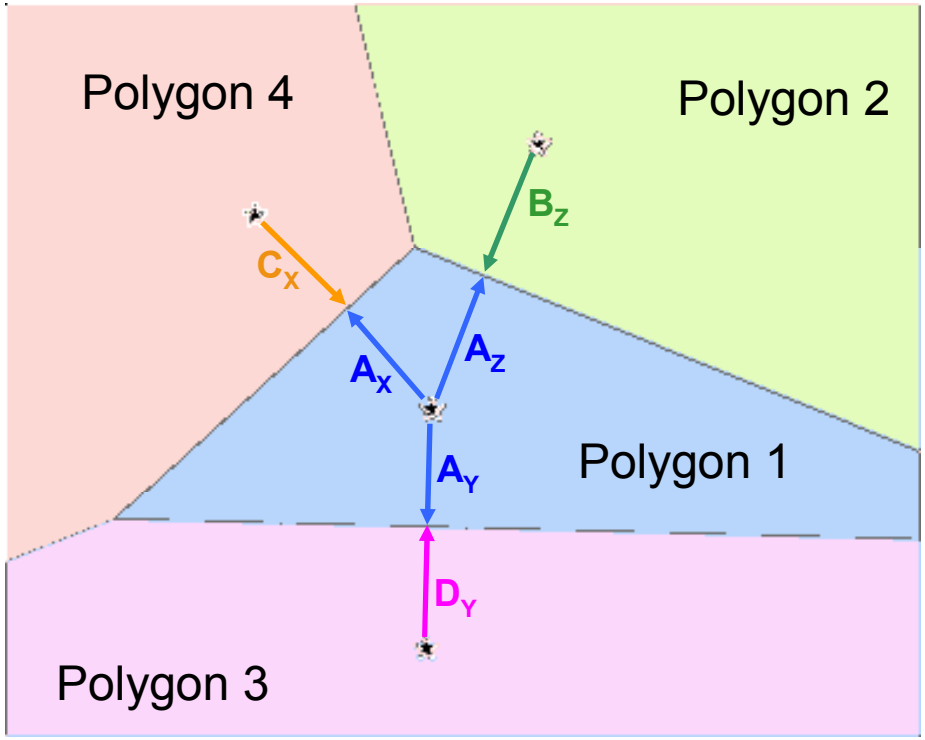
1. Create Theissen polygon coverage of monitoring sites
2. Link the 1990 and 2000 census tract polygons by tract ID to get total change in population by census tract
3. Convert census tract polygons to centroid points
4. Calculate the percent change in population for each monitoring area by spatially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



Map created in ESRI ArcMap 9.1

Aside: Thiessen polygons

- Thiessen polygons (also called Voronoi diagrams) are polygons whose boundaries define the area that is closest to each point relative to all other points.
- They are mathematically defined by the perpendicular bisectors of the lines between all points.



Thiessen Polygon Definitions

$A_X = C_X$

$A_Z = B_Z$

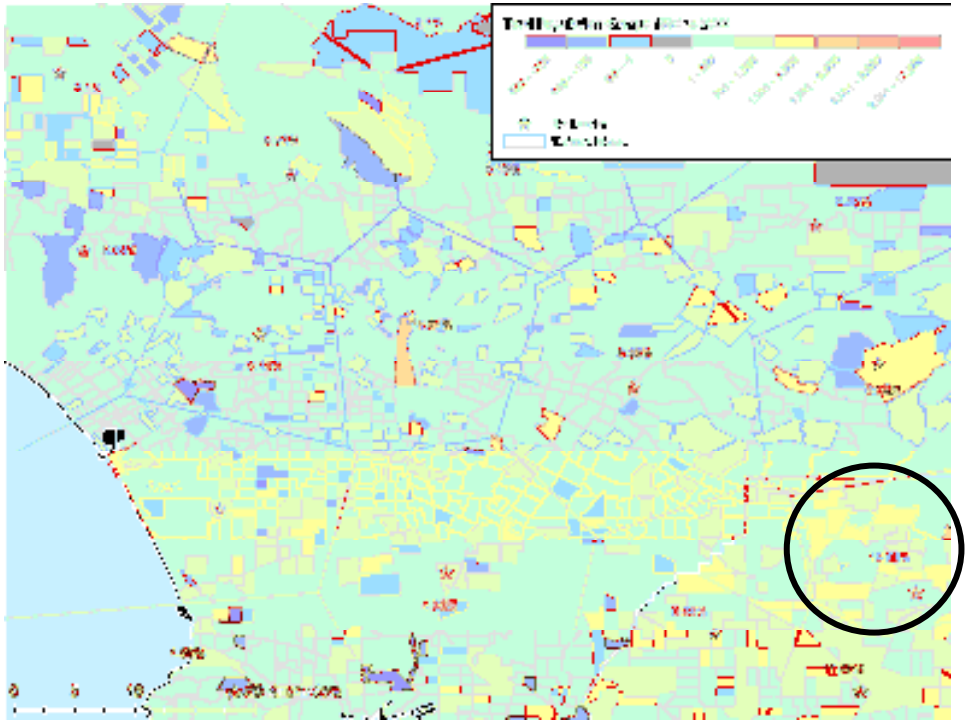
$A_Y = D_Y$

Black Stars: Monitor Locations

Population Change – Example

Interpretation

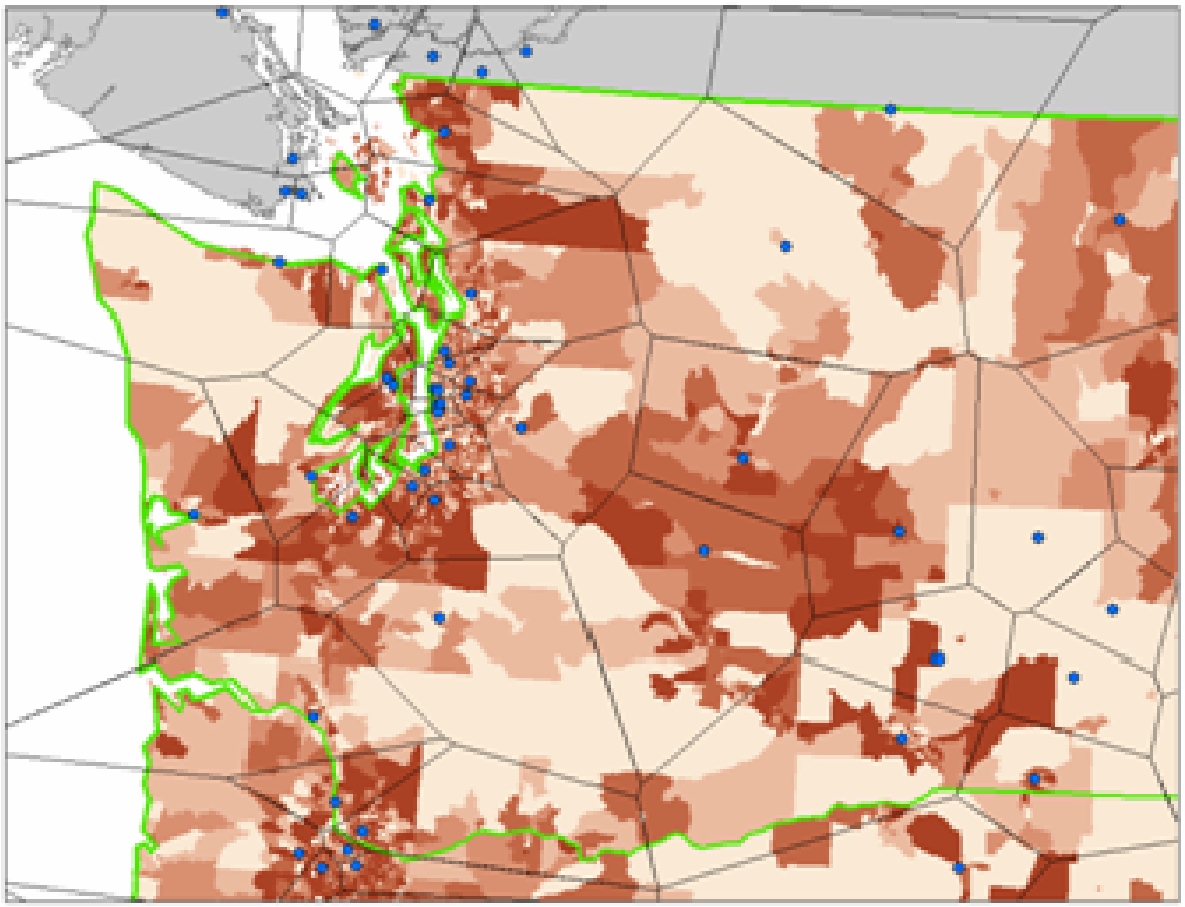
The area around site location 4 has seen a 13% increase in population and has, therefore, grown in importance for monitoring population exposure between 1990 and 2000.



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

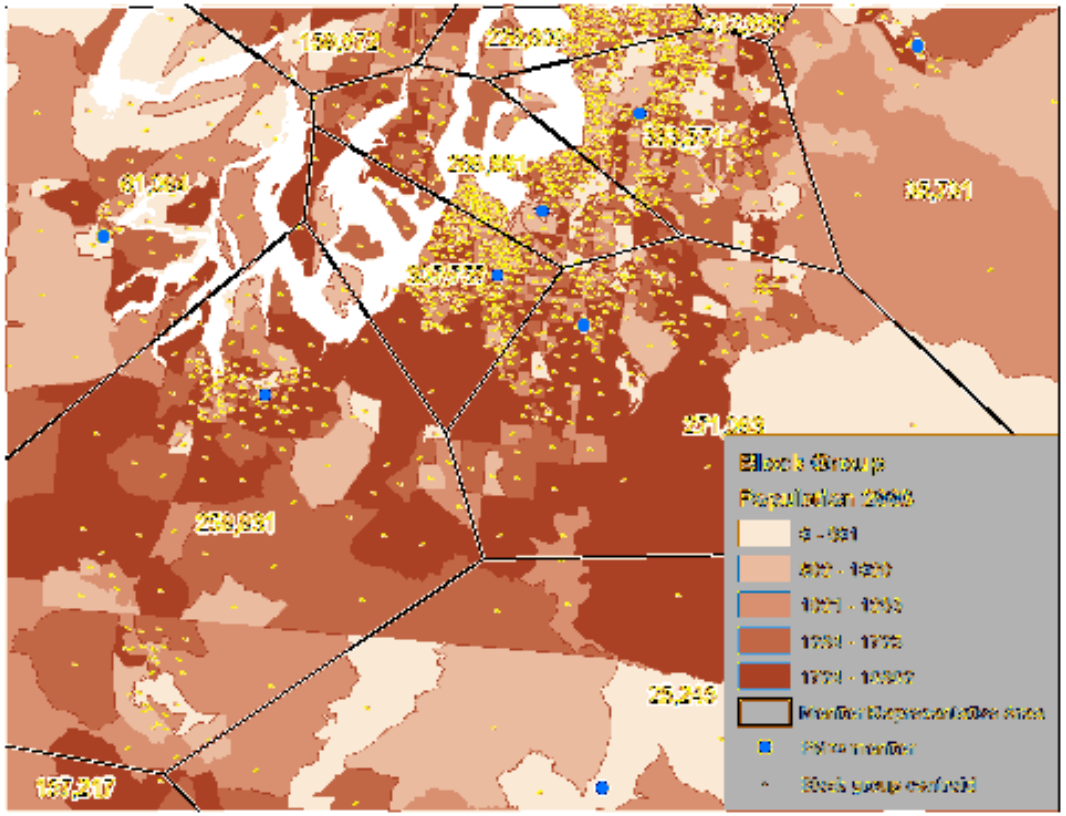
Population Served – Approach

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



Population Served – Example

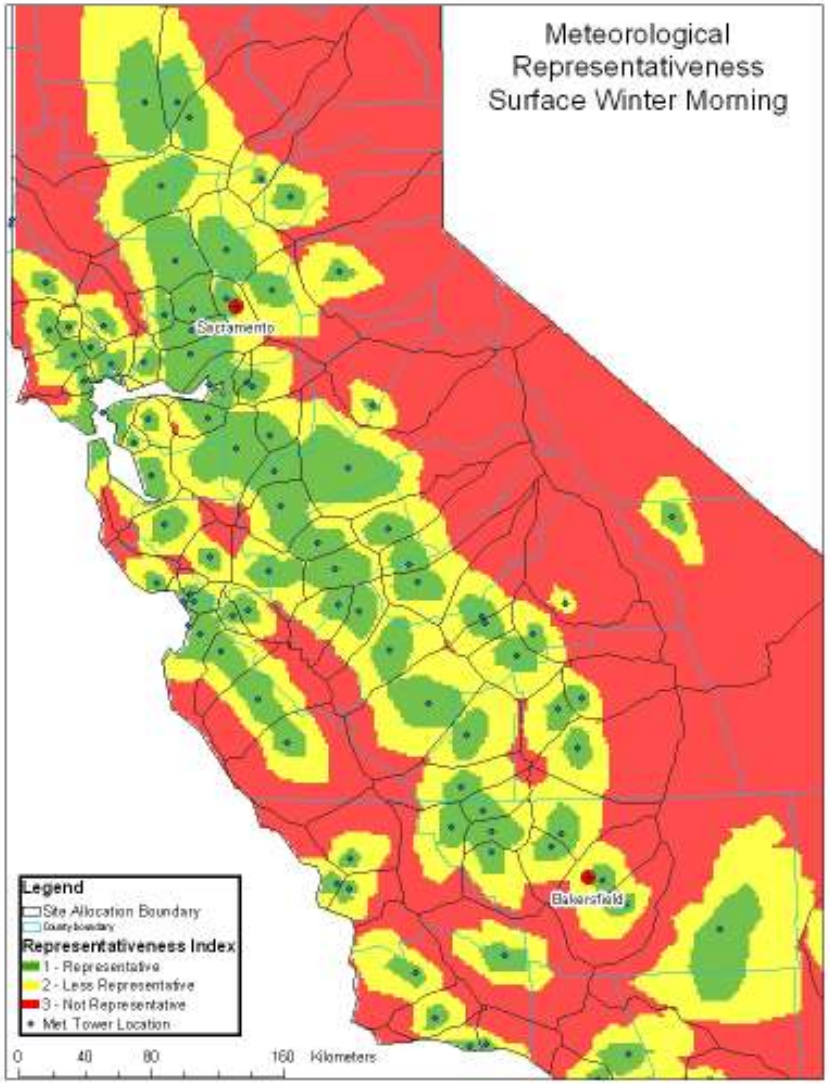
Note that the population served varies by two orders of magnitude. The actual population values could be used to *weight the sites*, or they could simply be ranked.



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

Area Served – Example

- More sophisticated techniques are available to determine areas of representation.
- This analysis considered meteorology, terrain, and distance within an empirical GIS model to determine areas well represented by meteorological monitoring towers.
- The closest monitor may not be the most representative of local conditions.



Removal Bias

Motivation:

- Removal bias is a sensitivity analysis to determine how important a particular monitor (or set of monitors) is for interpolating concentrations across the domain
- Measured values are interpolated across the domain using the entire network. Sites are then systematically removed and the interpolation is repeated

Resources needed:

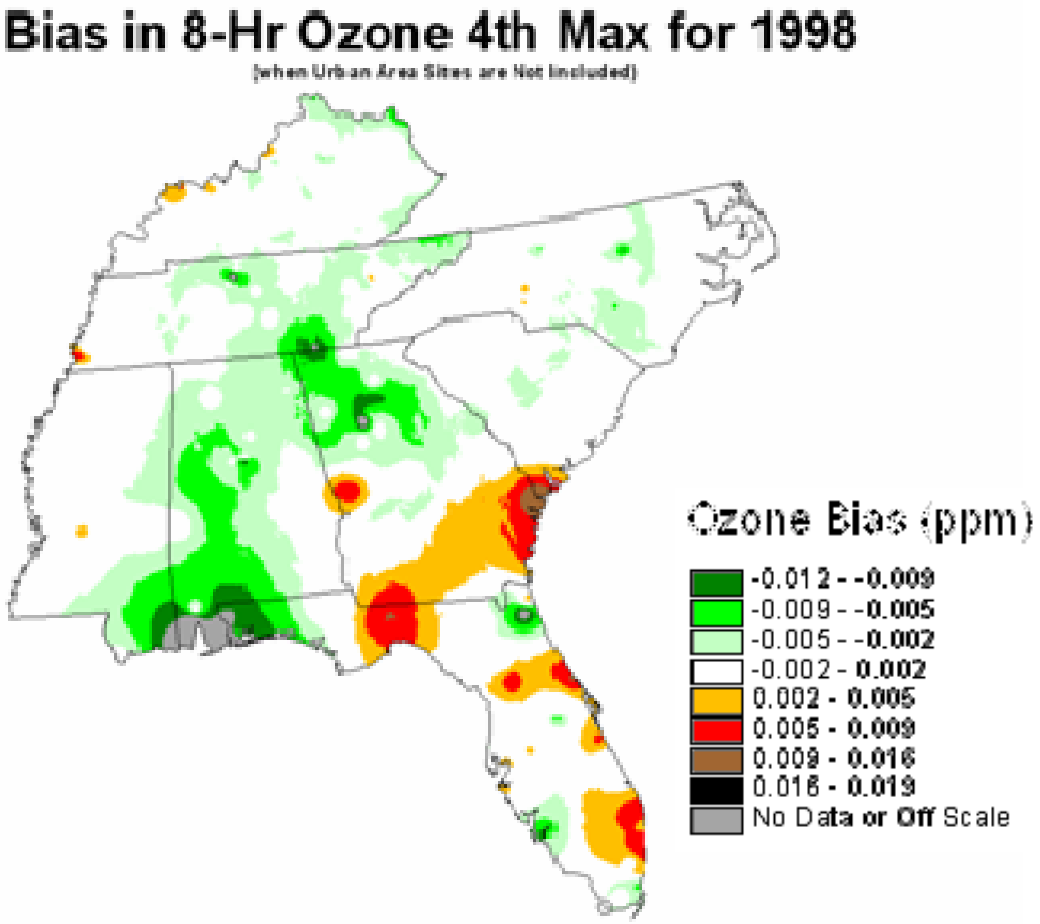
- Site location and concentration data from AQS or EPA AIRNow Tech
- GIS (geostatistical tools specifically)
- Statistical software may be helpful, but is not necessary

Removal Bias – Goals

- 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.
- 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.

Removal Bias – Example (1 of 2)

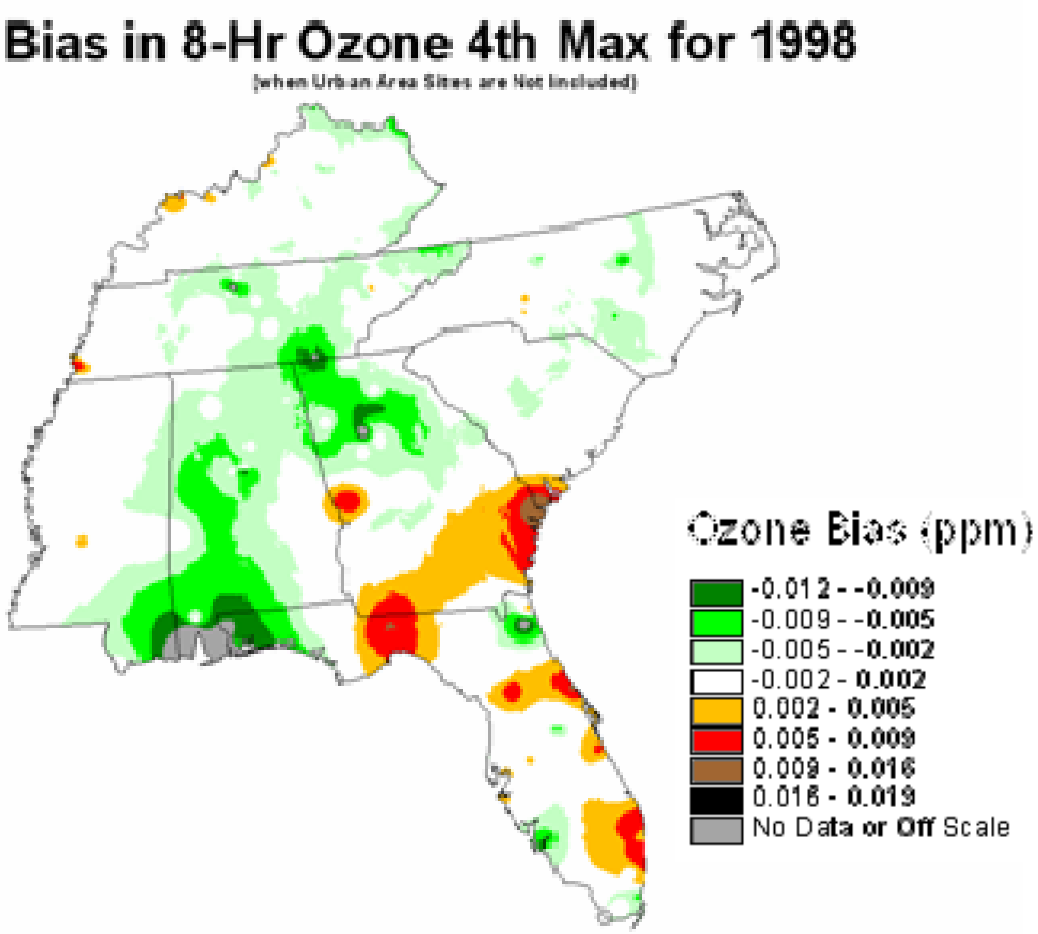
- This metric was one of five used in the 2000 National Analysis.
- Region 4 applied a network optimization technique, removing certain classes of sites (e.g., rural, urban core) and calculating interpolation bias.
- The map shows the bias in 8-hr ozone when all urban sites are removed - positive bias is shown in red and negative bias in green



Cimorelli et al, (2003) Region III ozone network reassessment.

Removal Bias – Example (2 of 2)

- Removing these urban monitors produces strong biases, both positive and negative.
- Negative bias from urban area site removal makes sense when maximum concentrations are often downwind, as with 8-hr ozone.
- This analysis can also be conducted by removing one site at a time. A large bias upon removal indicates a site contributing unique information.



Suitability Modeling

Motivation:

- Identifies suitable monitoring locations based on user-selected 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
- The results provide the best locations to site monitors based on the input criteria

Resources needed:

- GIS, site locations, population and other demographic/socioeconomic data, emission inventory data
- Meteorology and concentration data may be helpful, but are not necessary
- Skilled GIS analyst

Suitability Modeling – Example

Use GIS technology to

- Identify locations within an area where diesel particulate matter (DPM) emissions are likely to be high
- Identify locations potentially suitable for placing toxics and/or particulate monitors to better assess DPM impacts on population

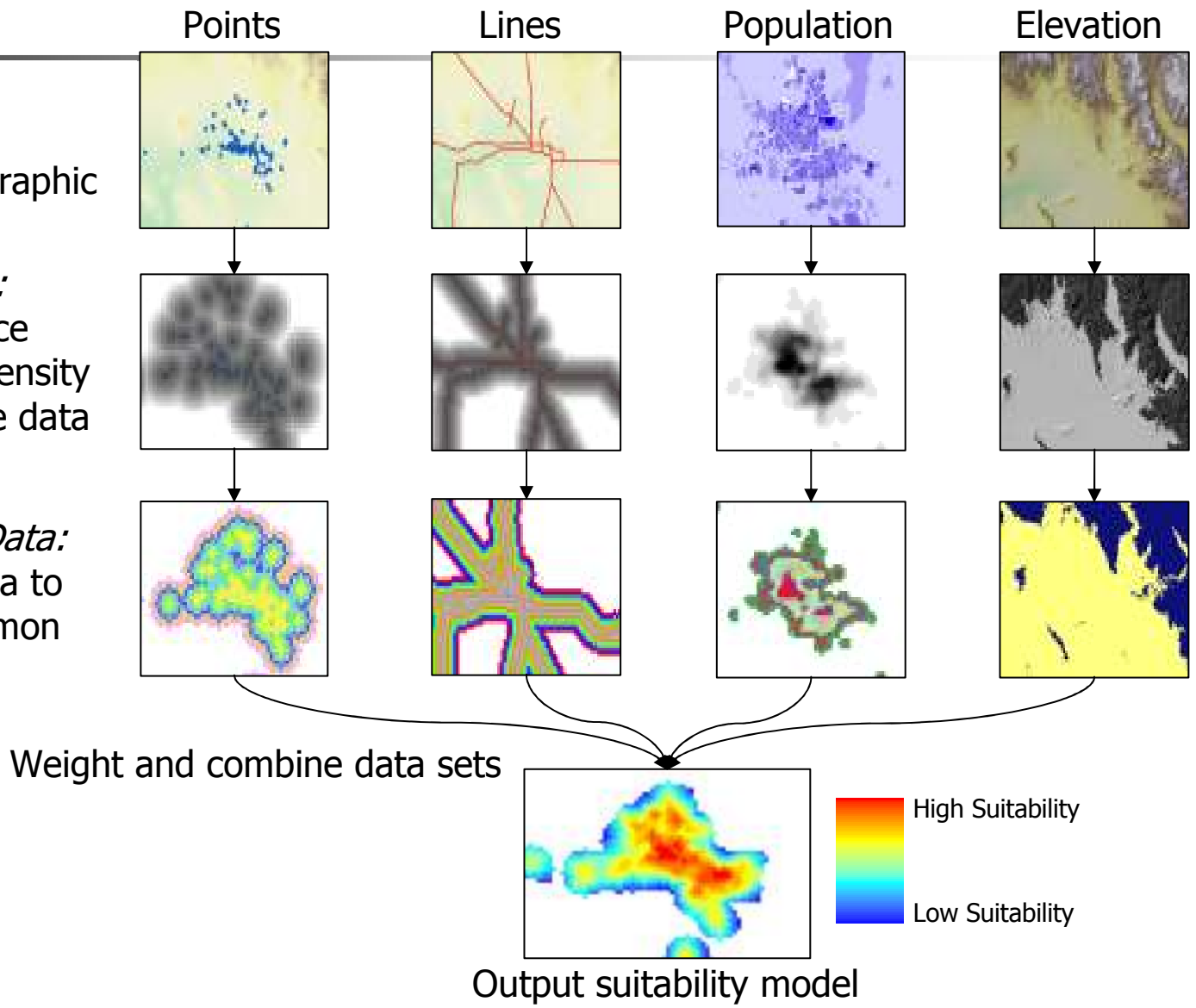
Training Example: Predicting Areas of High Diesel Particulate Matter Emissions in Phoenix, AZ, Using Spatial Analysis Techniques (2004)
Funded by the Arizona Department of Environmental Quality

Spatial Analysis Approach

Input Data:
Point, line, or polygon geographic data

Gridded Data:
Create distance contours or density plots from the data sets

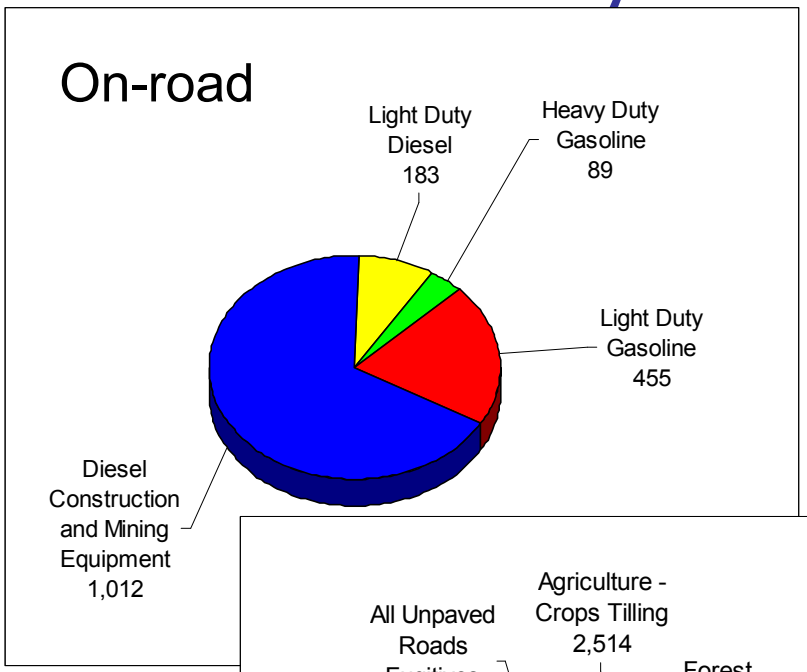
Reclassified Data:
Reclassify data to create a common scale

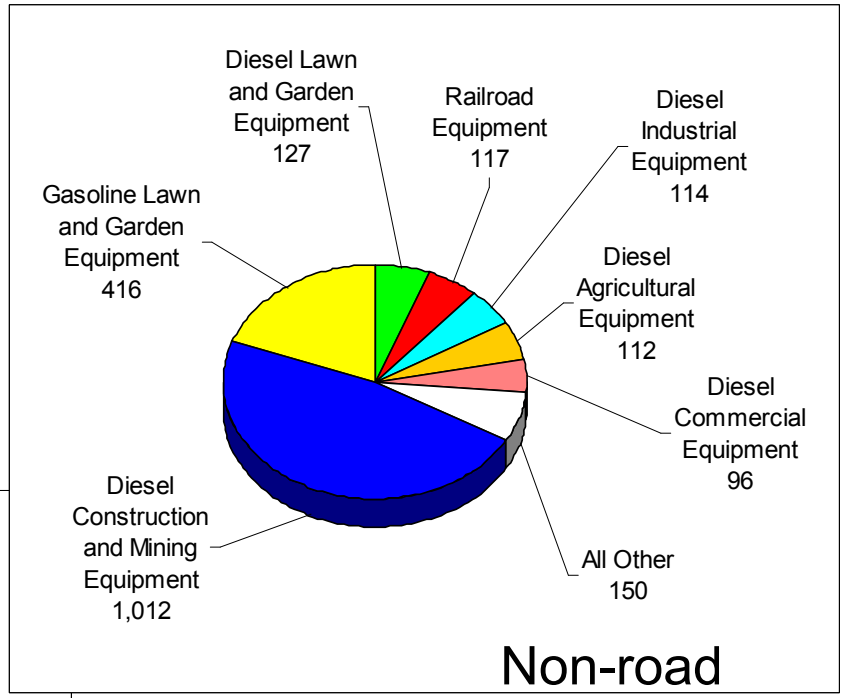


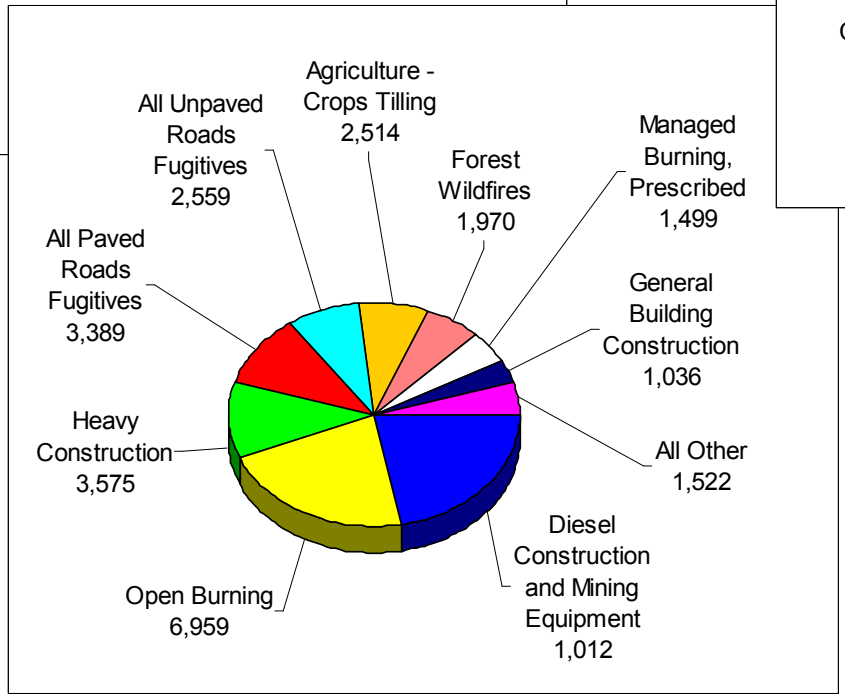
Analysis Method for DPM

- Assess the emission inventory to determine
  - Predominant sources of DPM
  - Best available geographic data to represent the spatial pattern of the identified emission sources in the region
- Determine the relative importance of each geographic data set based on its potential DPM contribution
- Weight input layers accordingly and combine the data sets to produce a suitability map using the GIS Spatial Analyst tool

Example of Using the Emission Inventory to Determine Layers



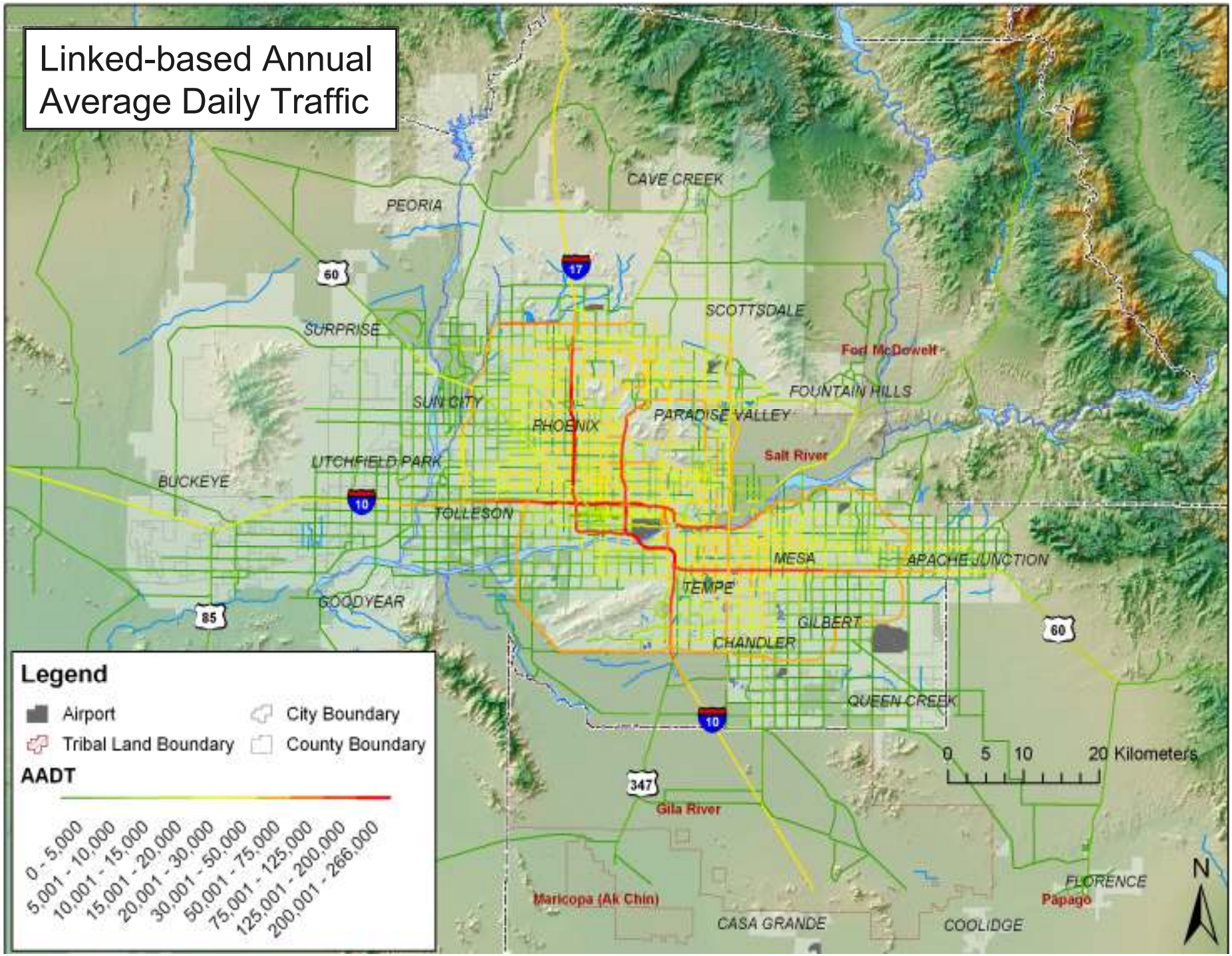


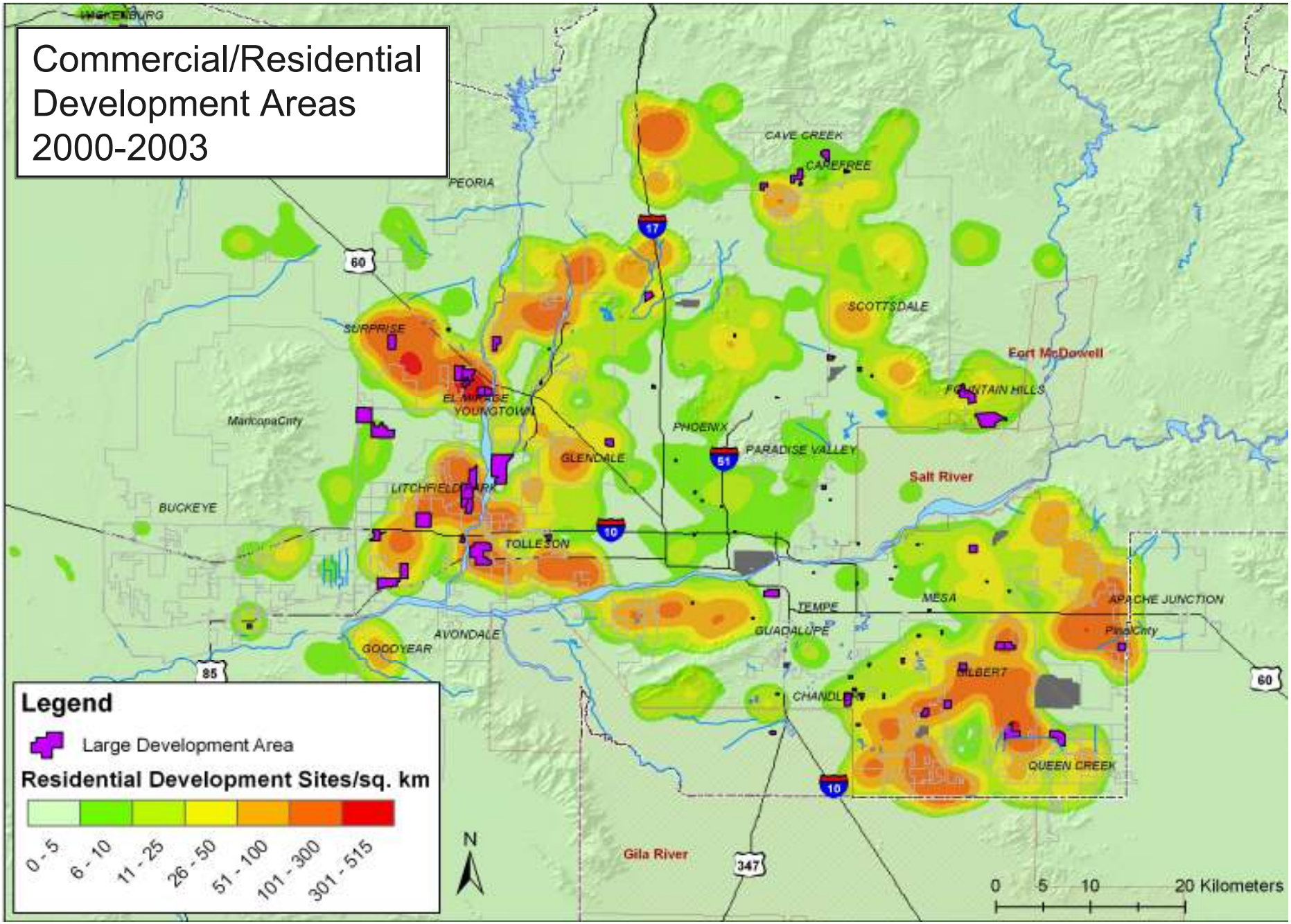


Major source category emissions of total $PM_{2.5}$ for Maricopa and Pinal Counties as reported in the 1999 National Emission Inventory

Data Layers Included

1. Traffic volume (Annual Average Daily Traffic, AADT)
2. Heavy-duty truck volume (from AADT data)
3. Locations of railroads and transportation depots
4. Residential and commercial development areas
5. Golf courses and cemetery locations (lawn and garden equipment usage)
6. Airport locations
7. $PM_{2.5}$ point source locations (weight assigned to each source depends on the source's relative EC contribution)
8. Total population and sensitive population (e.g., under 5 and over 65 years of age) density
9. Annual average gridded wind fields representing predominant wind direction throughout the region


Linked-based Annual Average Daily Traffic
CAVE CREEK
PEORIA
SURPRISE
SCOTTSDALE
Fort McDowell
FOUNTAIN HILLS
SUN CITY
PHOENIX
PARADISE VALLEY
Salt River
UTCHFIELD PARK
BUCKEYE
TOLLESON
MESA
APACHE JUNCTION
TEMPE
GOODYEAR
GILBERT
CHANDLER
QUEEN CREEK
Gila River
Maricopa (Ak Chin)
CASA GRANDE
COOLIDGE
FLORENCE
Papago
60
17
10
85
347
Legend
Airport
City Boundary
Tribal Land Boundary
County Boundary
AADT
0 - 5,000
5,001 - 10,000
10,001 - 15,000
15,001 - 20,000
20,001 - 30,000
30,001 - 50,000
50,001 - 75,000
75,001 - 125,000
125,001 - 200,000
200,001 - 266,000
0 5 10 20 Kilometers
N


Commercial/Residential Development Areas 2000-2003
CAVE CREEK
CAREFREE
PEORIA
SCOTTSDALE
SURPRISE
Fort McDowell
FOUNTAIN HILLS
EL MIRAGE
YOUNGTOWN
MaricopaCity
PHOENIX
PARADISE VALLEY
GLENDALE
Salt River
BUCKEYE
TOLLESON
TEMPE
MESA
APACHE JUNCTION
AVONDALE
GUADALUPE
GOODYEAR
GILBERT
CHANDLER
QUEEN CREEK
Gila River
Legend
Large Development Area
Residential Development Sites/sq. km
0 - 5
6 - 10
11 - 25
26 - 50
51 - 100
101 - 300
301 - 515
N
0 5 10 20 Kilometers

Phoenix Weighting Scheme

- Two model scenarios were used:
  1. Proximity to diesel emission sources (hot spot)
  2. Proximity of population to diesel sources
- Predominant wind direction was incorporated in each model scenario
  - For every point, direction to nearest feature in each layer was found (i.e., closest road)
  - Upwind, downwind, or cross-wind was defined for each point
  - Downwind influence was enhanced, but upwind influence was not subtracted
- Model scenario criteria were based on weighting assigned to each map layer depending on the layer's relative EC contribution

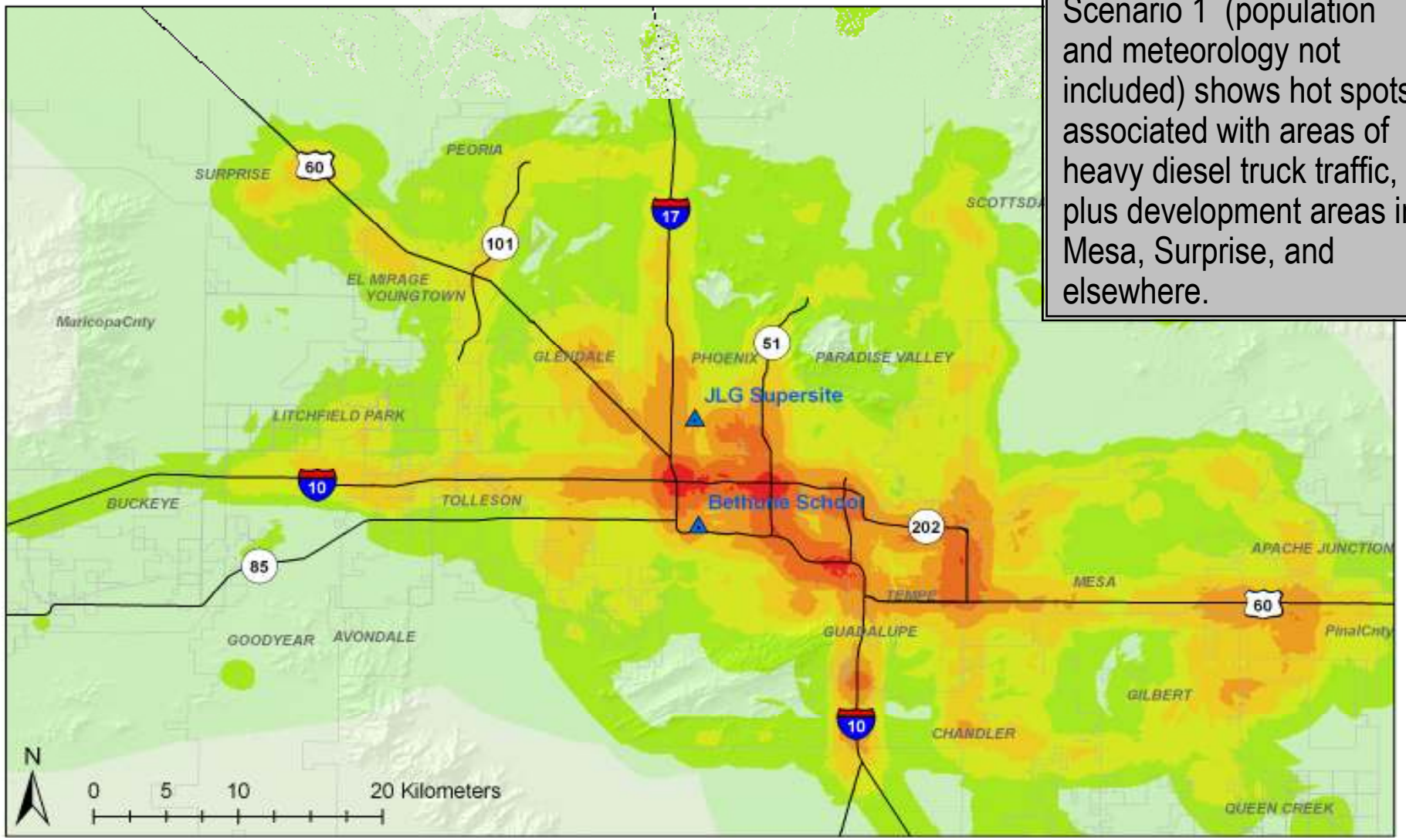
Phoenix Weighting Scheme

Two model scenarios were used:

1. Proximity to diesel emission sources (hot spot)
2. Proximity of population to diesel sources

Layer 1	(1) Hot Spot	(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 non-EC $PM_{2.5}$ emissions density = less suitable

Scenario 1 (population and meteorology not included) shows hot spots associated with areas of heavy diesel truck traffic, plus development areas in Mesa, Surprise, and elsewhere.



Legend

Suitability Model

PG Monitor Location

Interstate/Freeway

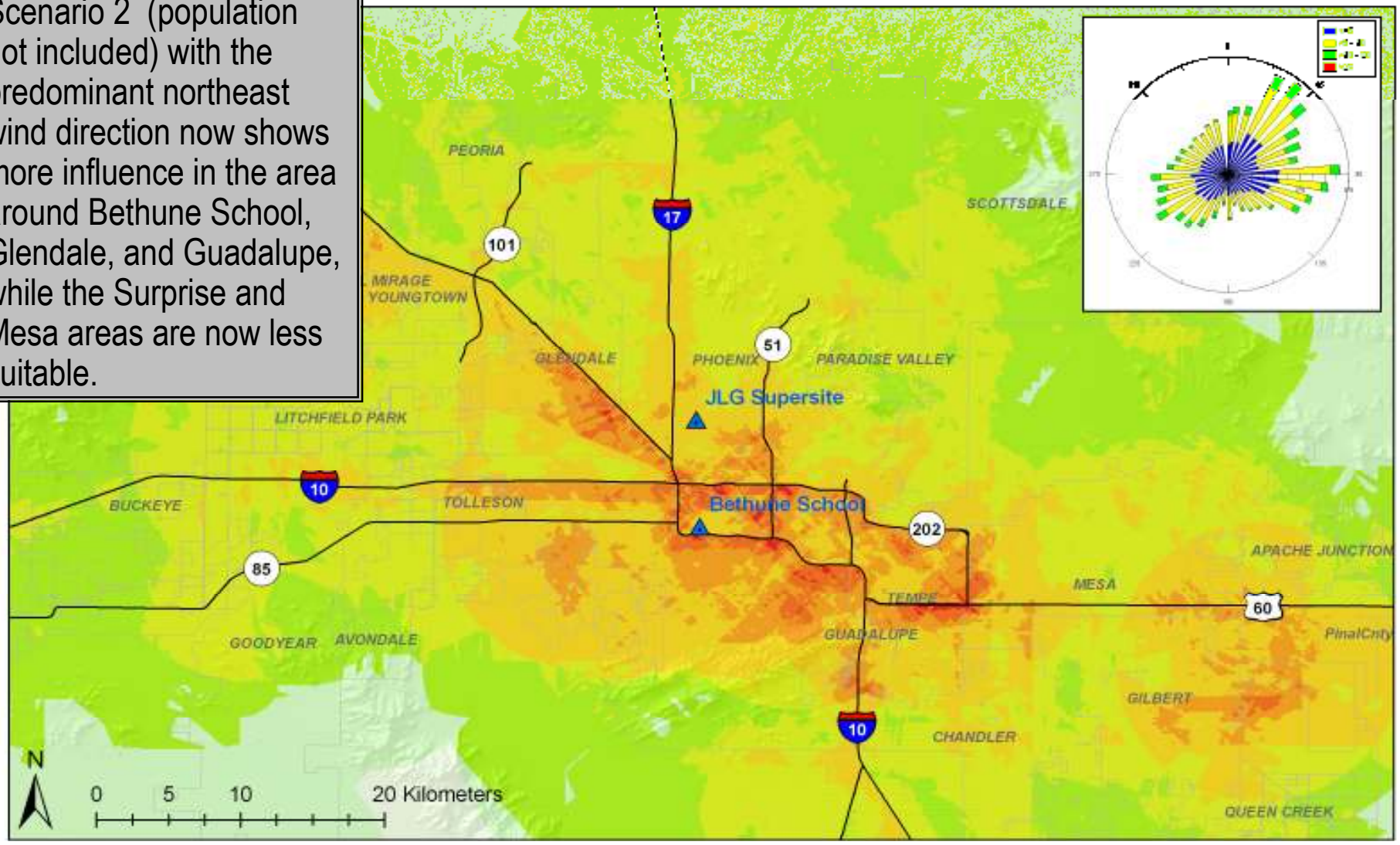
Urban Boundary

Hot Spot Weighting Scheme

Heavy Duty AADT Roads = 20%
Transportation Facilities = 20%
Commercial/Residential
Development Areas = 20%
Light Duty AADT Roads = 15%

Commercial Lawn/Garden
Cropp Areas = 10%
Rail & Road Corridors = 5%
Recreation = 5%
Airports = 5%

Scenario 2 (population not included) with the predominant northeast wind direction now shows more influence in the area around Bethune School, Glendale, and Guadalupe, while the Surprise and Mesa areas are now less suitable.



Legend

Suitability Model

PQ Monitor Location

Interstate/Freeway

Urban Boundary

Hot Spot/Wind Influence Weighting Scheme

Heavy Duty AADT Roads = 20%

Transportation Facilities = 20%

Commercial/Residential

Development Areas = 20%

Light Duty AADT Roads = 15%

Commercial Lawn/Garden

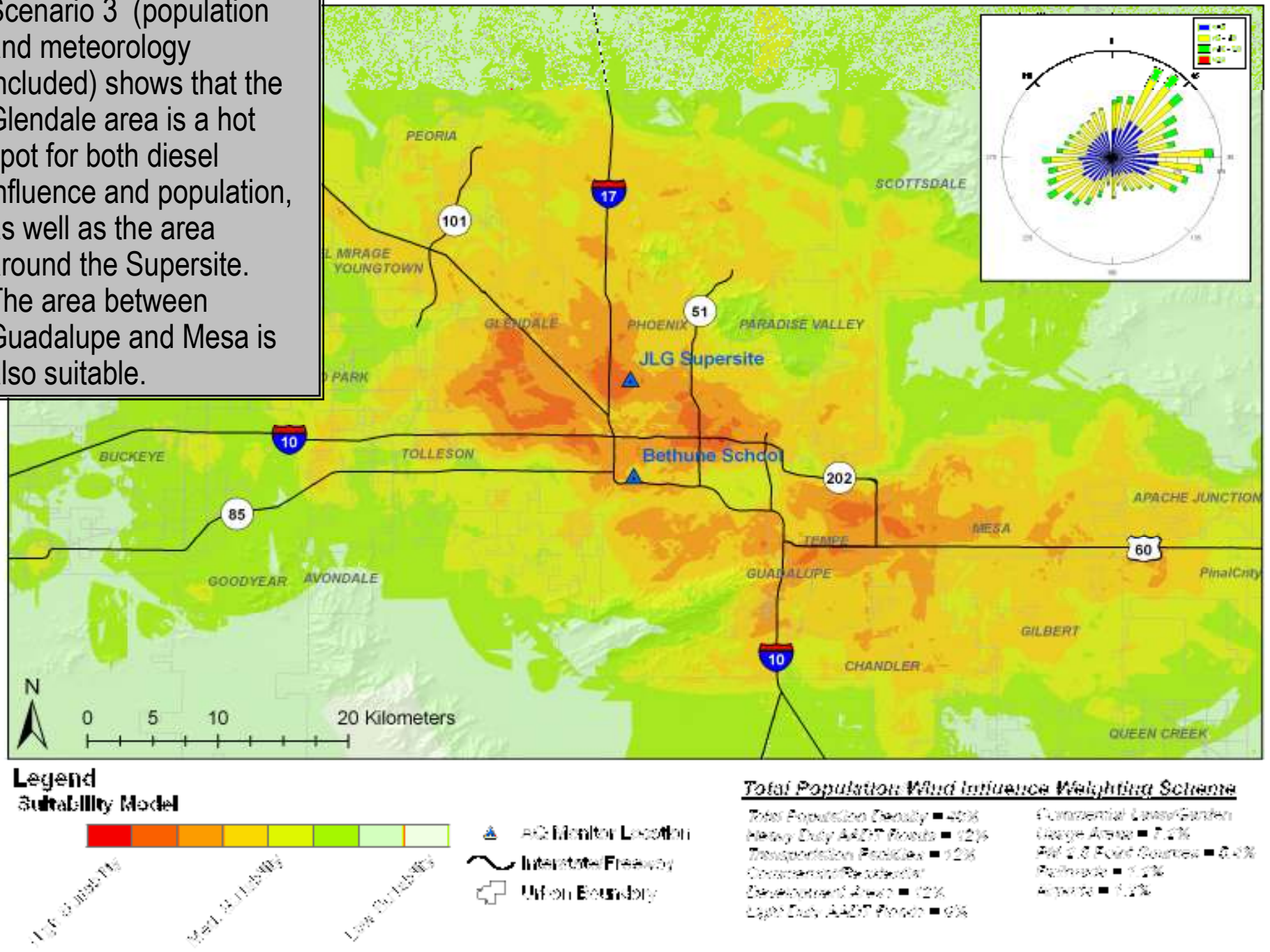
Large Areas = 10%

Pit & Point Sources = 5%

Railroads = 5%

Airports = 5%

Scenario 3 (population and meteorology included) shows that the Glendale area is a hot spot for both diesel influence and population, as well as the area around the Supersite. The area between Guadalupe and Mesa is also suitable.



Legend

Suitability Model

PG Monitor Location

Interstate Freeway

Urban Boundary

Total Population Wind Influence Weighting Scheme

Total Population Density = 40%
Heavy Duty AADT Roads = 12%
Transportation Facilities = 12%
Commercial/Recreation
Developed Area = 12%
Light Duty AADT Roads = 9%
Commercial Lawn/Garden
Urban Areas = 7.5%
Rail & Road Counties = 6.5%
Railroads = 1.2%
Airports = 1.2%

Suitability Analysis – Conclusions

- Results assist decision makers in
  - Assessing the utility of current monitors
  - Selecting locations for new monitors
  - Setting priorities for monitoring
  - Investigating a range of monitoring objectives and considerations
- Suitability analysis can improve the effectiveness of monitoring decisions

Principal Component Analysis (PCA)

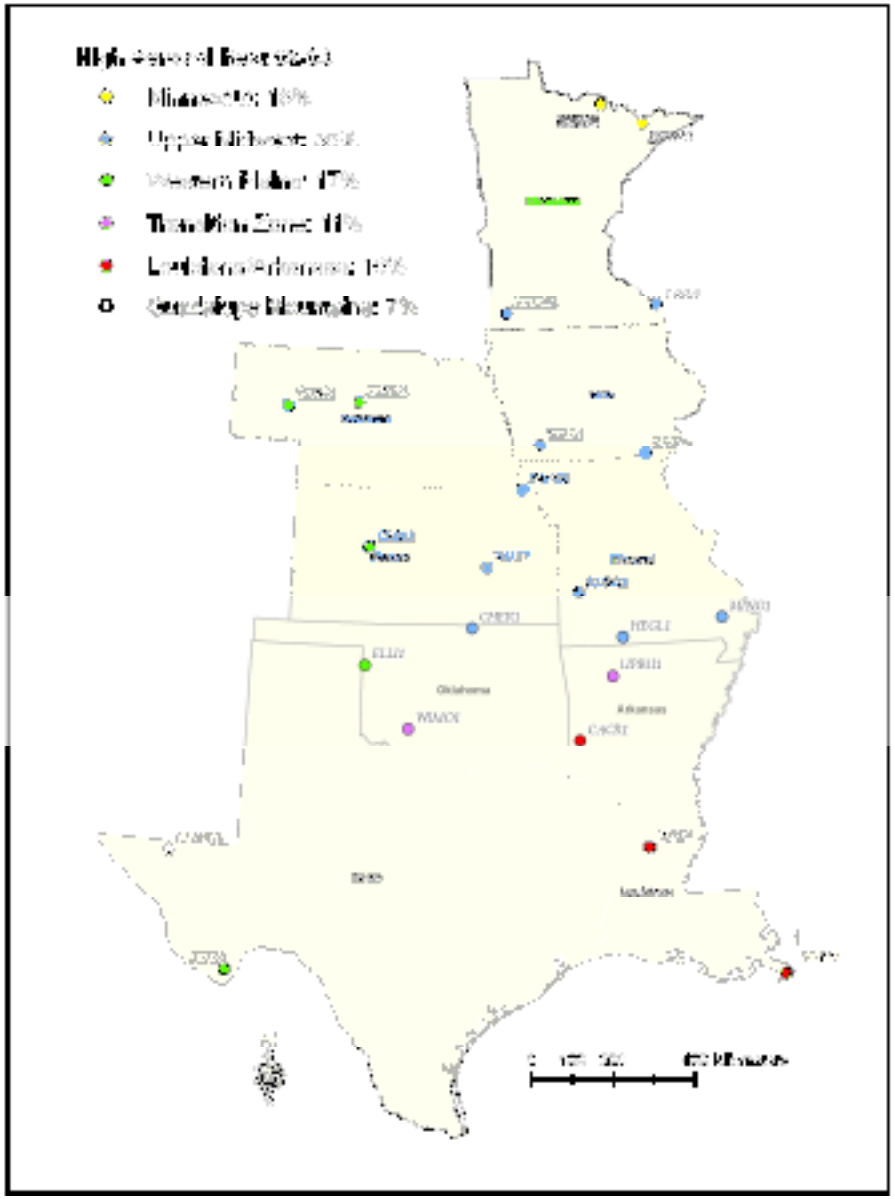
- Motivation: Find monitoring sites that have a pattern of variability similar to other monitoring sites
- Resources Needed:
  - Statistical software, concentration data, and site locations.
  - GIS and historical data would be helpful, but are not required

Training Example: Causes of Haze for the Central States Regional Air Planning Association (CENRAP) (2005)

Funded by the Central States Regional Air Planning Association (CENRAP)

Principle Component Analysis (PCA)

- Goal: Select representative sites for analysis by determining regions where aerosol extinction significantly covary in space and time.
- Sites in the same group (factor) have similar patterns and may be candidates for resource reallocation.
- Caution: Sites may covary while monitoring different magnitudes.



Positive Matrix Factorization (PMF)

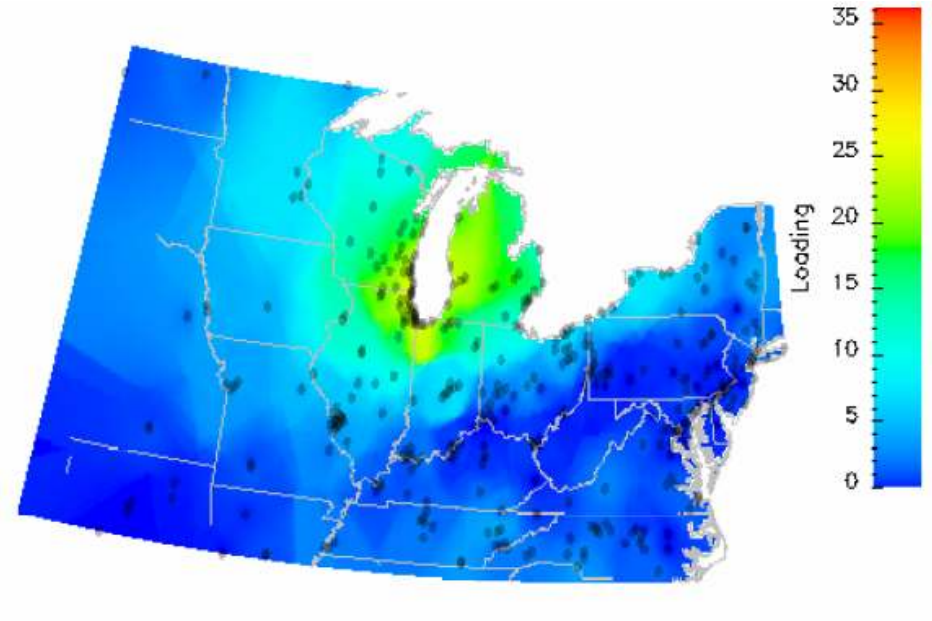
- Motivation:
  - Find monitoring sites that have a pattern of variability similar to other monitoring sites and assess the representativeness of individual sites
- Resources Needed:
  - Specialized software (i.e. EPA PMF, PMF2), concentration data, uncertainty estimates and site locations.
  - GIS, historical data and site info helpful though not required

Training Example: Assessing Ozone Networks Using Positive Matrix Factorization (Rizzo and Scheff) (2004)

Funded by Region 5, Environmental Protection Agency (EPA)

Positive Matrix Factorization (PMF) Example – EPA Region 5 O_3 Networks

- Goal: Group sites into regions of similar variability and identify specific monitors to be removed or relocated.
- Sites in the same group (factor) have similar patterns and may be redundant. In addition, PMF predicts concentrations; ratios of actual to predicted concentrations can be used to select specific sites that are or are not contributing useful information about ozone concentrations in the region.



Example of one factor group for ozone monitoring sites in EPA region 5 (USEPA, 2003).

Additional PMF Resources

- EPA Multivariate Receptor Modeling Workbook
  - http://www.epa.gov/heasd/products/pmf/pmf.htm
  - Contact Shelly Eberly at eberly.shelly@epa.gov

Conclusions

- Networks must be assessed to ensure that the considerable resources required to run the networks are used optimally.
- Monitoring networks can fulfill many scientific, regulatory, and outreach objectives.
- Assessment of network (and individual site) efficacy depends on objectives.
- A wide-range of analytical techniques of varying complexity can address these objectives.

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.

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).

Knoderer C.A. and Raffuse S.M. (2004) CRPAQS surface and aloft meteorological representativeness (California Regional PM10/PM2.5 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).

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.