UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
Office of Air Quality Planning and Standards
Research Triangle Park, North Carolina 27711

August 29, 2003

MEMORANDUM

SUBJECT: Analyses of 2000-2002 PM Data for the PM NAAQS Review

FROM: Mark Schmidt, OAQPS
David Mintz, OAQPS
Tesh Rao, OAQPS
Lance McCluney, OAQPS (Intern)

TO: file

The purpose of this memorandum is to describe and summarize multiple sets of analyses conducted for the review of the Particulate Matter (PM) national ambient air quality standards (NAAQS). Both PM_{10} and $PM_{2.5}$ data were analyzed, as well as the calculated differences of the two particle size cuts ($PM_{10-2.5}$), and $PM_{2.5}$ composition data. Most PM_{10} and $PM_{2.5}$ data, and corresponding meteorological information, were extracted from EPA's Air Quality System (AQS) data base on various dates in May, 2003. $PM_{2.5}$ composition data from urban sites in the EPA Speciation Network (ESpN) were pulled from AQS in July, 2003. PM mass and $PM_{2.5}$ composition data from rural sites in the Interagency Monitoring of PROtected Visual Environmental (IMPROVE) aerosol monitoring network were acquired from the National Park Service in July, 2003. Additional meteorological data were obtained from the National Weather Service. Meteorological data were necessary to convert AQS PM_{10} samples reported at 'standard conditions' (25° C, 760 mm Hg) to 'local conditions' (actual temperature and pressure). The conversion was necessary to facilitate size cut comparisons of PM_{10} to $PM_{2.5}$ and calculate an accurate difference; PM_{10} data are generally reported to AQS at standard conditions and $PM_{2.5}$ data are reported at local conditions. There are three attachments to this memo, each corresponding to the different type of data analyzed: Attachment A describes the AQS-based, 24-hour duration analyses; Attachment B describes the AQS-based hourly analyses; and Attachment C describes the ESpN and IMPROVE data analyses. Each attachment contains a discussion of the methods and assumptions used to generate results. All AQS-based 24-hour duration PM (10 and 2.5 micron size cuts) data and hourly PM_{10} data used in the analyses were sampled with Federal Reference Methods (FRM) or Federal Equivalent Methods (FEM). Hourly AQS $PM_{2.5}$ data and particle data collected in the ESpN and IMPROVE networks (Attachment C) utilized non-FRM/FEM techniques.

States are required to certify their data submitted to AQS on an annual basis for each calender year; this certification must be done by July 1st of the following year. Since the 2002 data used for these analyses were queried from AQS prior to the certification deadline, it should

be noted that the 2002 data are subject to change and that additional 2002 data may have been reported after the retrievals used in these analyses.

Some analysis results are summarized at a broad regional level using the geographic regions specified below. The regional definitions correspond to the regions identified by the Health Effects Institute (HEI) in a recent PM study. [See Figure 1, page 8, in Samet, J.M., et al., "The National Morbidity, Mortality, and Air Pollution Study Part II: Morbidity, Mortality, and Air Pollution in the United States," Health Effects Institute, Research Report Number 94, Part II, June 2000.] The origin of the HEI region definitions can be traced back to Figure 6-30 of EPA's 1996 PM Criteria Document, which identified regions on the basis of "uniqueness in aerosol trends, seasonality, size distribution, or chemical composition." Some sites (e.g., ones in Alaska, Hawaii, Puerto Rico, and the Virgin Islands,) were not assigned to an HEI region. For these analyses, these sites were placed in Region 0, 'Not in PM Region'. Data for these sites are excluded from charts shown 'by region' but are included elsewhere.

PM REGION CODE	PM REGION DESCRIPTION	HOW DEFINED
1	Northeast	ME, NH, VT, MA, RI, CT, NJ, DE, MD*, PA*, NY*, VA*, WV* (*east of -78.50° W longitude)
2	Southeast	NC, SC, TN, GA, FL, AL, MS, LA, AR, OK*, TX* (*east of -97.70° W longitude)
3	Industrial Midwest	NY*, PA*, WV*, VA*, KY, OH, MI, IN, IL, WI#, MN#, IA#, MO# (*west of -78.50° W longitude, #east of -91.50° W longitude)
4	Upper Midwest	MN*, WI*, IA*, MO*, ND, SD, NE, KS, CO# (*west of -91.50° W longitude, #east of -104.05° W longitude)
5	Southwest	OK*, TX*, NM, AZ, NV#, CA# (*west of -97.70° W longitude, #south of 37.00° N latitude and east of -115.50° W longitude)
6	Northwest	WA, ID, MT, WY, UT, OR, CO*, CA#, NV# (*west of -104.05° W longitude, #north of 37.00° N latitude)
7	Southern California	CA*, NV* (*west of -115.50° W longitude and south of 37.00° N latitude)

For additional information on the analyses documented in the attachments, please contact Mark Schmidt at (919) 541-2416.

3 Attachments

ATTACHMENT A

Processing Details for AQS 24-hour Sample Duration Files and Figures

This attachment describes the data (2000-2002) and processing procedures used to generate the following AQS-based, 24-hour sample duration files and figures:

- File PM25_sitemon_info.xls: $PM_{2.5}$ Site-Monitor Information
- File PM10_sitemon_info.xls: PM_{10} Site-Monitor Information
- File PMC_sitemon_info.xls: $PM_{10-2.5}$ Site-Monitor Information
- File PM25_sitemon_summary.xls: $PM_{2.5}$ Monitor Data Summary
- File PM10_sitemon_summary.xls: PM_{10} Monitor Data Summary
- File PMC_sitemon_summary.xls: $PM_{10-2.5}$ Monitor Data Summary
- File PM25ctymax.xls: $PM_{2.5}$ County Max Data Summary
- File PM10ctymax.xls: PM_{10} County Max Data Summary
- File PMCctymax.xls: $PM_{10-2.5}$ County Max Data Summary
- Figure 2-4. Distribution of annual mean $PM_{2.5}$ and estimated annual mean $PM_{10-2.5}$ concentrations by region, 2000-2002.
- Figure 2-5. Distribution of 98th percentile 24-hour average $PM_{2.5}$ and estimated $PM_{10-2.5}$ concentrations by region, 2000-2002.
- Figure 2-6. County-level maximum annual mean $PM_{2.5}$ concentrations, 2000-2002.
- Figure 2-7. County-level maximum 98th percentile 24-hour average $PM_{2.5}$ concentrations, 2000-2002.
- Figure 2-9. County-level maximum annual mean PM_{10} concentrations, 2000-2002.
- Figure 2-10. County-level maximum 98th percentile 24-hour average PM_{10} concentrations, 2000-2002.
- Figure 2-11. Estimated county-level maximum annual mean $PM_{10-2.5}$ concentrations, 2000-2002.
- Figure 2-12. Estimated county-level maximum 98th percentile 24-hour average $PM_{10-2.5}$ concentrations, 2000-2002.
- Figure 2-16. Distribution of ratios of annual mean $PM_{2.5}$ to PM_{10} by region, 2000-2002.
- Figure 2-17. Regional average correlation of 24-hour average PM by size fraction.
- Figure 2-18. Urban 24-hour average $PM_{2.5}$ concentration distributions by region and month, 2000-2002.
- Figure 2-19. Urban 24-hour average $PM_{10-2.5}$ concentration distributions by region and month, 2000-2002.
- Figure 2-20. Distribution of annual mean vs. 98th percentile 24-hour average $PM_{2.5}$ concentrations, 2000-2002.
- Figure 2-21. Distribution of estimated annual mean vs. 98th percentile 24-hour average $PM_{10-2.5}$ concentrations, 2000-2002.

General Data Description

All data, except for supplemental meteorological data obtained from NWS, were extracted from EPA's Air Quality System database (AQS). After downloading and necessary preprocessing, data for PM_{10}, $PM_{2.5}$, and calculated $PM_{10-2.5}$, sites were subjected to data completeness criteria. The data selection criteria for all PM size cuts (applied independently to PM_{10}, $PM_{2.5}$, and calculated $PM_{10-2.5}$) was (by site) the most recent 4, 8, or 12 consecutive quarters of 11 or more samples. A simple example is shown below. For this example site, the quarters that would have been utilized are shaded. Since the selection criterion evaluates available data in increments of 4 quarters, previous quarters could not be used due to the shortfall in 2001, quarter 1. An additional increment of 4 consecutive quarters meets the 11 minimum sample threshold (1999, quarters 1-4), but would not have been used since the more recent band of data (shaded) were available. Although the utilized selection criteria do not guarantee a calendar year(s) of data, it does provide at least one full year consisting of four quarters, thus reducing seasonal bias. Data present in quarters not part of the 4-, 8-, or 12-quarter period of interest were deleted and thus, not included in a site summaries

	'00 Q1	'00 Q2	'00 Q3	'00 Q4	'01 Q1	'01 Q2	'01 Q3	'00 Q4	'00 Q4	'02 Q1	'02 Q3	'02 Q4
N=	12	13	14	15	10	15	16	14	15	13	11	9

Means and percentiles were calculated for each site that met completeness criteria. Weighted 'annual' means (referenced as 'ANNMEAN' in summary data files) were computed for each site as follows: quarterly averages were calculated for each kept quarter; 4-quarter averages were then computed from the applicable one, two, or three sets of quarterly averages (e.g., in the example above, from the '00 Q4, '01 Q1, '01 Q2, and '01 Q3 averages); then, the 4-quarter means were averaged together. Percentiles of interest (minimum, maximum, median, 5^{th}, 25^{th}, 75^{th}, and 95^{th}) were computed on the entire 4-, 8-, or 12-quarter distribution of data.

Although all data submitted to AQS are considered valid, some data are tagged with quality assurance qualifiers, natural event flags, and/or exceptional event flags. All data, regardless of flags, were included in these analyses.

All concentrations presented in the analyses outputs are shown in units of micrograms per cubic meter ($\mu g/m^3$) at local conditions. Since most regulatory PM_{10} data are reported in units of $\mu g/m^3$ at standard temperature and pressure conditions (25° C, 760 mm Hg), the data had to be converted to local temperature and pressure conditions using meteorological information (see next section). The standard conditions PM_{10} data were converted to local conditions before the site completeness criteria (most recent 4, 8, or 12 consecutive quarters of 11+ samples) were applied.

PM_{10} Data

PM_{10} data from Federal Reference Methods (FRM) and Federal Equivalent Methods (FEM) monitors were extracted from AQS on May 28, 2003. Four separate queries were made: 1) raw daily (24-hour) for parameter 81102 [PM_{10}, standard temperature and pressure conditions (STP)]; 2) raw daily for parameter 85101 [PM_{10}, local temperature and pressure conditions (LTP)]; 3) summary daily (hourly reported measurements aggregated within AQS to a 24-hour period) for parameter 81102; and 4) summary daily for parameter 85101. The 'daily' monitors collected 24-hour average samples on a filter for each successful day of monitoring. The monitors are typically scheduled to collect PM_{10} samples once every 6 days, though some collected samples more frequently. The PM_{10} filter samples are weighed in a laboratory environment to obtain mass concentrations expressed in micrograms per cubic meter ($\mu g/m^3$). The 'hourly' monitors are generally operated continuously almost every hour of the year (with occasional down time for calibrations and audits). AQS maintains the raw hourly data and also aggregates the hourly information into summary daily records. A summary record is only created if 75% or more of the hourly data ($\geq$16) are present.

Parameter 81102 data, both summary and daily, were converted to local conditions using collocated temperature and pressure information. If collocated temperature and/or pressure data were not available, meteorological data from the nearest NWS station were used. If collocated data were not available and the NWS data were missing, the STP data were not converted to LTP and not used in the analyses. Parameter 85101 and converted 81102 data were merged by site day. If more than one type of PM_{10} data were present on a given day, LTP reported data were favored over converted (STP to LTP) data, and lower Pollutant Occurrence Codes (POC's) were favored over higher POC numbers. Typically, monitors with lower POC numbers will have more data (sample more frequently) than monitors with higher POC numbers (for the same duration data). 'Primary' sampler POC numbers are generally lower than quality assurance (QA) monitor POC's; thus primary data were favored over QA data. Daily sampling monitors

also usually have lower POC numbers than hourly sampling instruments, so the daily data were generally selected over the summary daily data. Hence, the following selection priority was implemented: 85101 daily > 85101 summary > 81102 daily > 81102 hourly.

Site data were evaluated on the 11 sample per quarter threshold and a determination made as to which 4, 8, or 12 (all) quarters to keep. 1235 sites had at least 4 consecutive quarters of 11+ samples: 707 sites had 12 quarters available, 246 sites had 8 quarters available, and 282 sites had 4 quarters to use. The 1235 sites are mapped in Figure B-1. Metadata for the 1235 sites are provided in file, 'PM_{10}_sitemon_info.xls'. Summary data for the sites (means and moments) are provided in file, 'PM_{10}_sitemon_summary.xls'. County maxima of site 'ANNMEAN' and 98th percentile are provided in 'PM10ctymax.xls'.

$PM_{2.5}$ Data

$PM_{2.5}$ data from non-continuous FRM monitors were extracted from AQS on May 28, 2003. Only one AQS query was necessary: 24-hour (daily) data for parameter 88101 ($PM_{2.5}$, local temperature and pressure conditions). Hourly data were not extracted for these analyses since are there are no continuous FRM/FEM methods for $PM_{2.5}$. $PM_{2.5}$ monitors are typically scheduled to collect $PM_{2.5}$ samples once every 3 days, though some collected samples more or less frequently. The $PM_{2.5}$ filter samples are weighed in a laboratory environment to obtain mass concentrations ($\mu g/m^3$). $PM_{2.5}$ data, reported as parameter 88101, are in local conditions. Only primary monitors were used in the analyses; primary monitors are the first occurring POC, generally '1'. Site data were evaluated on the 11 sample per quarter threshold and a determination made as to which 4, 8, or 12 (all) quarters to keep. 1152 sites had at least 4 consecutive quarters of 11+ samples: 789 sites had all 12 quarters available, 193 sites had 8 quarters available, and 170 sites had 4 quarters to use. The 1152 sites are mapped in 'pm25-map.gif'. Metadata for the 1152 sites are provided in file, 'PM25_sitemon_info.xls'. Summary data for the sites (means and moments) are provided in file, 'PM25_sitemon_summary.xls'. County maxima (of site 'ANNMEAN' and 98th percentile) are provided in 'PM25ctymax.xls'.

$PM_{(10-2.5)}$ Estimates

In order to characterize a PM coarse fraction (i.e., PM less than 10 micrometers but greater than 2.5 micrometers), a simplistic difference method was utilized. At locations where both PM_{10} and $PM_{2.5}$ were recorded, $PM_{2.5}$ daily averages are subtracted from PM_{10} daily averages. Currently there are no federal PM coarse fraction monitoring requirements. Although there are no federal reference or equivalent methods stipulated for the PM coarse fraction, only FRM/FEM PM_{10} and $PM_{2.5}$ daily averages were used to construct the $PM_{10-2.5}$ estimates. No effort was made to account for differences in sampling instruments or protocols between the co-located PM_{10} and $PM_{2.5}$ monitors. Because of these differences (and other factors), occasionally the calculated $PM_{10-2.5}$ values were negative; this is not unexpected for two independent observations and negative $PM_{10-2.5}$ concentrations were not censored from the analyses. Both the PM_{10} and $PM_{2.5}$ data used in the difference calculation were in units of $\mu g/m^3$ at local conditions, thus the calculated $PM_{10-2.5}$ values also are in those units. All available FRM/FEM PM_{10} and $PM_{2.5}$ data were used to construct the $PM_{10-2.5}$ estimates; the completeness criteria (most recent 4, 8, or 12 consecutive quarters of 11+ samples) were evaluated *after* the daily difference estimates were calculated. 488 sites met the completeness criteria of at least 4 consecutive quarters of 11+

samples; 219 sites had 12 quarters available, 129 sites had 8 quarters available, and 140 sites had 4 quarters to use. The 488 sites are mapped in Figure B-3. Metadata for the 488 sites are provided in file, 'PMC_sitemon_info.xls'. Summary data for the sites (means and moments) are provided in file, 'PMC_sitemon_summary.xls'. County maxima (of site 'ANNMEAN' and 98th percentile) are provided in 'PMCctymax.xls'.

Boxplot Figures

Many of the analyses figures are boxplots. Unless otherwise noted, in all of the AQS-based, 24-hour average duration boxplots, the following definitions apply:

- The bottom of the box depicts the 25th percentile of the plotted distribution
- The top of the box depicts the 75th percentile of the plotted distribution
- The line through the box identifies the distribution median
- The top whisker cap identifies the 95th percentile of the plotted distribution
- The bottom whisker cap identifies the 5th percentile of the plotted distribution
- The distribution maximum and minimum are shown as asterisks

Data Files and Processing Code

The graphics generated for the PM NAAQS review and the 3 size faction site maps are included in the attached file, 'Attach-A-Graphics'.

The nine Microsoft Excel spreadsheets mentioned above are contained in the attached file, 'Attach-B-spreadsheets.zip' along with a data dictionary. A spreadsheet of boxplot plotting points is also included.

Raw, intermediate, and final SAS data files generated for these analyses are included in the attached file, 'Attach-B-SASdata.zip':

Data were processed with SAS software. The SAS programs are included in the attached file, 'Attach-B-SAS.zip':

Comments on Specific Figures:

Figure 2-4. Distribution of annual mean $PM_{2.5}$ and estimated annual mean $PM_{10-2.5}$ concentrations by region, 2000-2002:

- Shows distribution of site level ANNMEAN by size fraction and region.
- Region 0 data (23 sites for $PM_{2.5}$, 13 sites for $PM_{10-2.5}$) were excluded from chart.

Figure 2-5. Distribution of 98th percentile 24-hour average $PM_{2.5}$ and estimated $PM_{10-2.5}$ concentrations by region, 2000-2002:

- Shows distribution of site level PCT98 by size fraction and region.
- Region 0 data (23 sites for $PM_{2.5}$, 13 sites for $PM_{10-2.5}$) were excluded from chart.

Figure 2-6. County-level maximum annual mean $PM_{2.5}$ concentrations, 2000-2002:

- Shows county maximum of $PM_{2.5}$ site mean (referenced as ANNMEANMAX in

PM25ctymax.xls)
- Concentration ranges provided by AQSSD, HEEG.

Figure 2-7. County-level maximum 98th percentile 24-hour average $PM_{2.5}$ concentrations, 2000-2002:
- Shows county maximum of $PM_{2.5}$ site 98th percentile (referenced as PCT98MAX in PM25ctymax.xls)
- Concentration ranges provided by AQSSD, HEEG.

Figure 2-9. County-level maximum annual mean PM_{10} concentrations, 2000-2002:
- Shows county maximum of PM_{10} site mean (referenced as ANNMEANMAX in PM10ctymax.xls)
- Concentration ranges provided by AQSSD, HEEG.

Figure 2-10. County-level maximum 98th percentile 24-hour average PM_{10} concentrations, 2000-2002:
- Shows county maximum of PM_{10} site 98th percentile (referenced as PCT98MAX in PM_{10}ctymax.xls)
- Concentration ranges provided by AQSSD, HEEG.

Figure 2-11. Estimated county-level maximum annual mean $PM_{10-2.5}$ concentrations, 2000-2002:
- Shows county maximum of $PM_{10-2.5}$ site mean (referenced as ANNMEANMAX in PMCctymax.xls)
- Concentration ranges provided by AQSSD, HEEG.

Figure 2-12. Estimated county-level maximum 98th percentile 24-hour average $PM_{10-2.5}$ concentrations, 2000-2002:
- Shows county maximum of $PM_{10-2.5}$ site 98th percentile (referenced as PCT98MAX in PMCctymax.xls)
- Concentration ranges provided by AQSSD, HEEG.

Figure 2-16. Distribution of ratios of 24-hour average $PM_{2.5}$ to PM_{10} by region, 2000-2002:
- The ratio of $PM_{2.5}$ to PM_{10} was first calculated for each site-day. Because the parameter selection criteria (most recent 4, 8, or 12 consecutive quarters of 11+ samples) were applied separately for PM_{10} and $PM_{2.5}$, the selected time periods did not necessarily match. If the common time periods of constituent raw data (for the PM10 and PM2.5 sites that met the selection criteria) were used for this analysis, some sites common to both parameters would not have any matches (by site-day) and others would have a seasonal bias (only have matches in certain quarters). To avoid this situation, the raw data used in this analysis were culled from the $PM_{10-2.5}$ database (pmc_raw_meetscomp.sas7bdat). This insured an equal number of each quarter for each site and also insured a minimum of 44 samples for each site (4 quarters * 11 samples each).
- The site-day ratios of $PM_{2.5}$ to PM_{10} were averaged by site and the distribution of the site ratios plotted by region.

Figure 2-17. Regional average correlation of 24-hour average PM by size fraction:
- For the same reason noted in the first bullet above, all data used in this analysis were culled from the $PM_{10-2.5}$ database (PMC_raw_meetscomp.sas7bdat).
- A Pearson correlation coefficient was calculated for each site fraction pair (PM_{10} versus $PM_{2.5}$, $PM_{2.5}$ versus $PM_{10-2.5}$, and PM_{10} versus $PM_{10-2.5}$.
- The site correlation coefficients for each fraction were averaged by region.
- Region 0 data (23 sites for $PM_{2.5}$, 13 sites for $PM_{10-2.5}$) were excluded from the chart.

Figure 2-18. Urban 24-hour average $PM_{2.5}$ concentration distributions by region and month,

2000-2002.

- Only data from monitors with location setting (referenced 'LOCATION" in 'PM25_sitemon_info.xls') = 'URBAN AND CENTER CITY' or 'SUBURBAN' were used.
- All 24-hour average values from kept quarters at above noted monitors were averaged together by region-month.
- In the plots, the boxes represent the interquartile range (25^{th} to 75^{th} percentiles) of each monthly distribution and the line inside the box is the median of the distribution. The trend line represents the mean, and the number above each box represents the number of 24-hour average observations that were used to generate each box plot.
- Only valid regions (1-7) were plotted.
- Seven separate graphics were produced for the PM NAAQS review memo; a 2-character HEI region name abreviation differentiates the attached plots.

Figure 2-19. Urban 24-hour average $PM_{10-2.5}$ concentration distributions by region and month, 2000-2002.

- Only data from monitors with location setting (referenced 'LOCATION" in 'PMC_sitemon_info.xls') = 'URBAN AND CENTER CITY' or 'SUBURBAN' were used.
- All 24-hour average values from kept quarters at above noted monitors were averaged together by region-month.
- In the plots, the boxes represent the interquartile range (25^{th} to 75^{th} percentiles) of each monthly distribution and the line inside the box is the median of the distribution. The trend line represents the mean, and the number above each box represents the number of 24-hour average observations that were used to generate each box plot.
- Only valid regions (1-7) were plotted.
- Seven separate graphics were produced for the PM NAAQS review memo; a 2-character HEI region name abbreviation differentiates the attached plots.

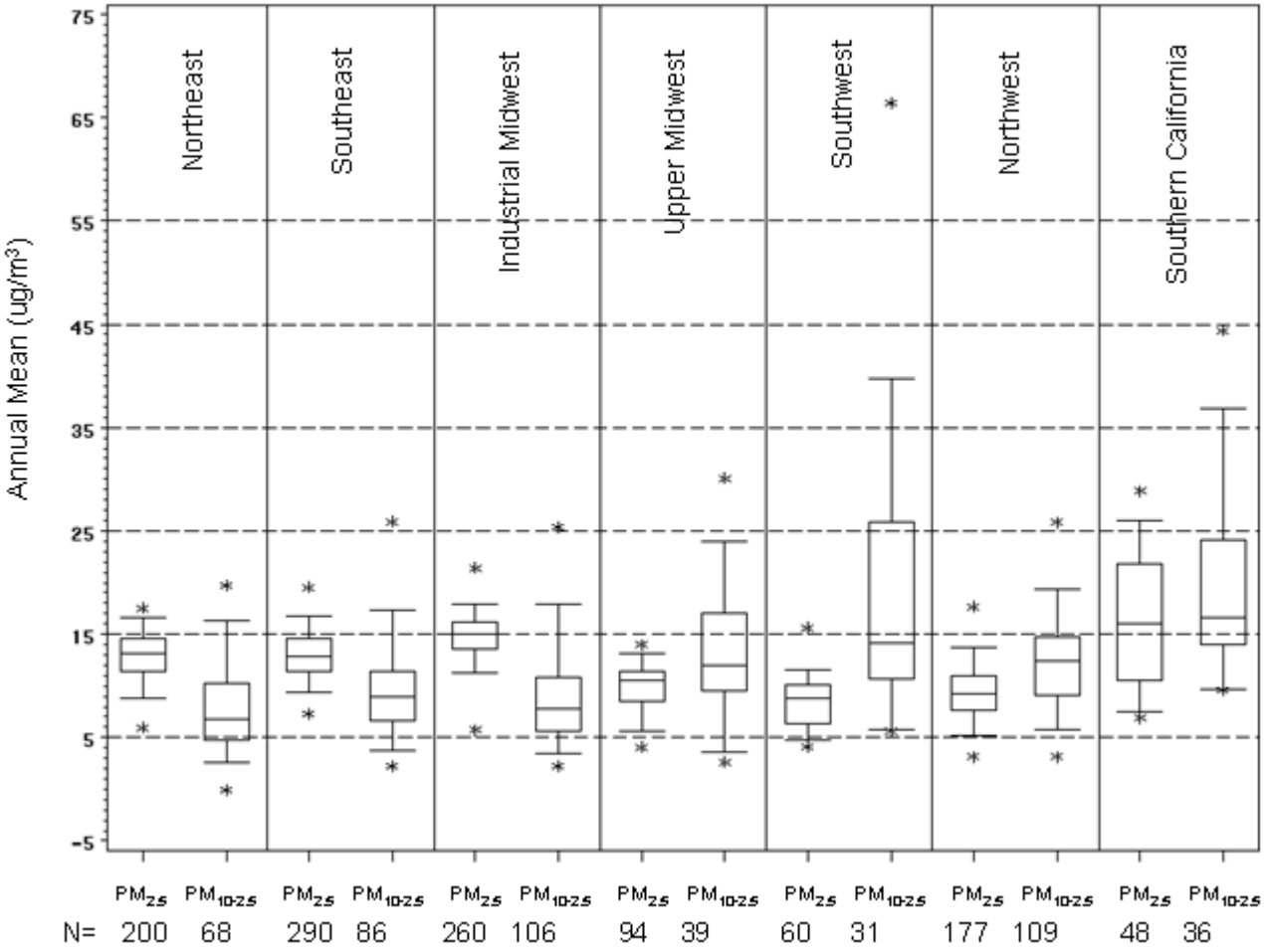
Figure 2-20. Distribution of annual mean vs. 98^{th} percentile 24-hour average $PM_{2.5}$ concentrations, 2000-2002:

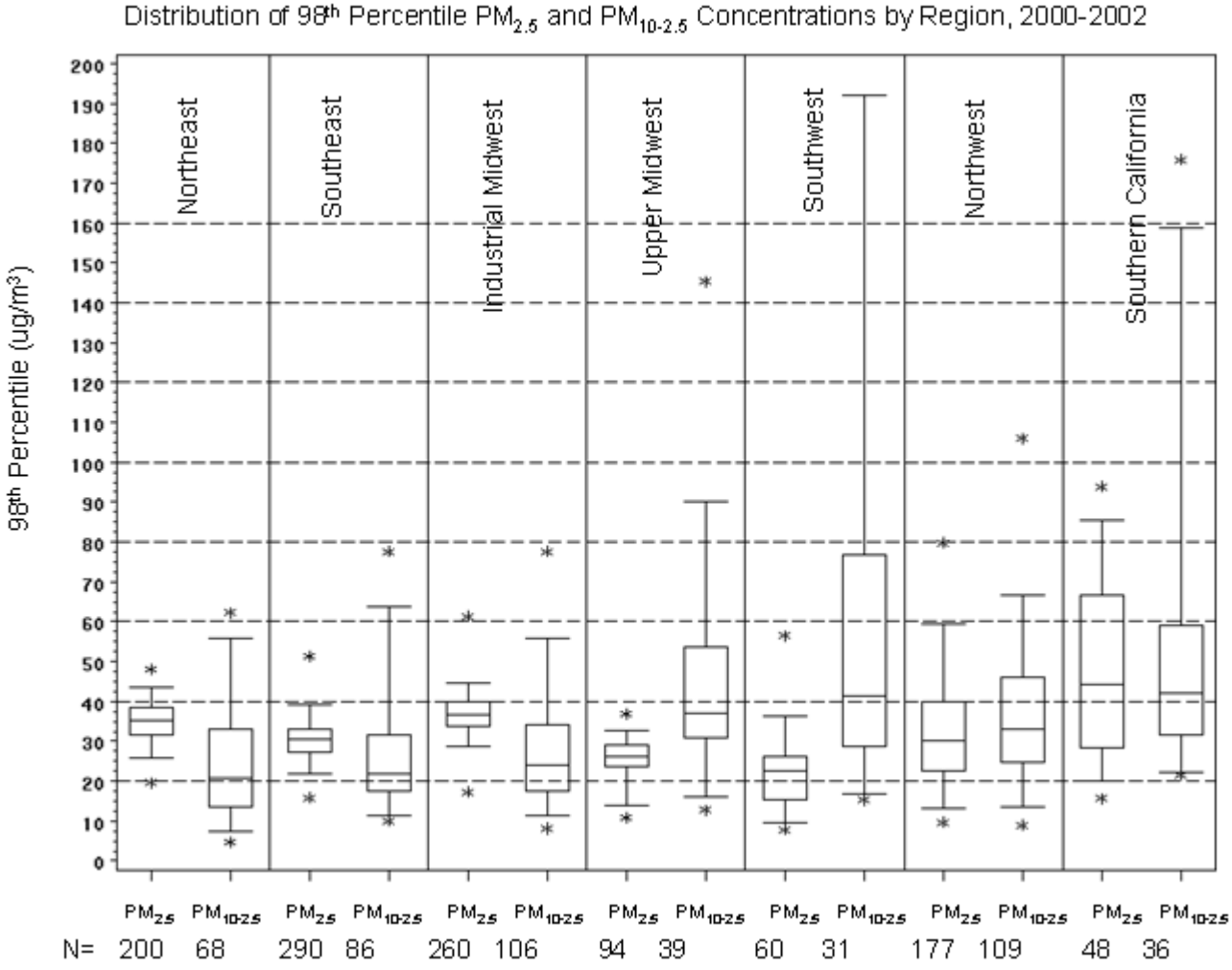
- Shows distribution of site level 98^{th} percentile by annual mean range.
- Concentration ranges provided by AQSSD, HEEG.

Figure 2-21. Distribution of estimated annual mean vs. 98^{th} percentile 24-hour average $PM_{10-2.5}$ concentrations, 2000-2002:

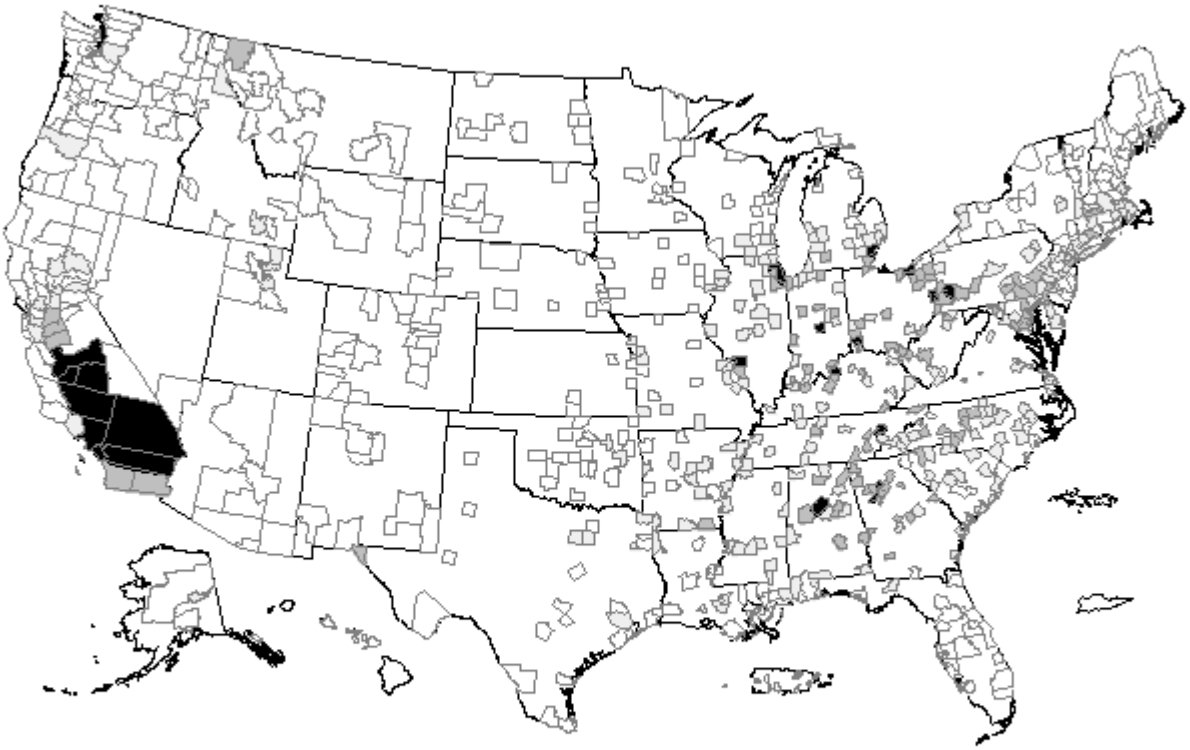
- Shows distribution of site level 98^{th} percentile by annual mean range.
- Concentration ranges provided by AQSSD, HEEG.

Distribution of Annual Mean $PM_{2.5}$ and $PM_{10-2.5}$ Concentrations by Region, 2000-2002




Distribution of 98th Percentile PM2.5 and PM10-2.5 Concentrations by Region, 2000-2002
98th Percentile (ug/m3)
Northeast
Southeast
Industrial Midwest
Upper Midwest
Southwest
Northwest
Southern California
PM2.5
PM10-2.5
N= 200 68 290 86 260 106 94 39 60 31 177 109 48 36

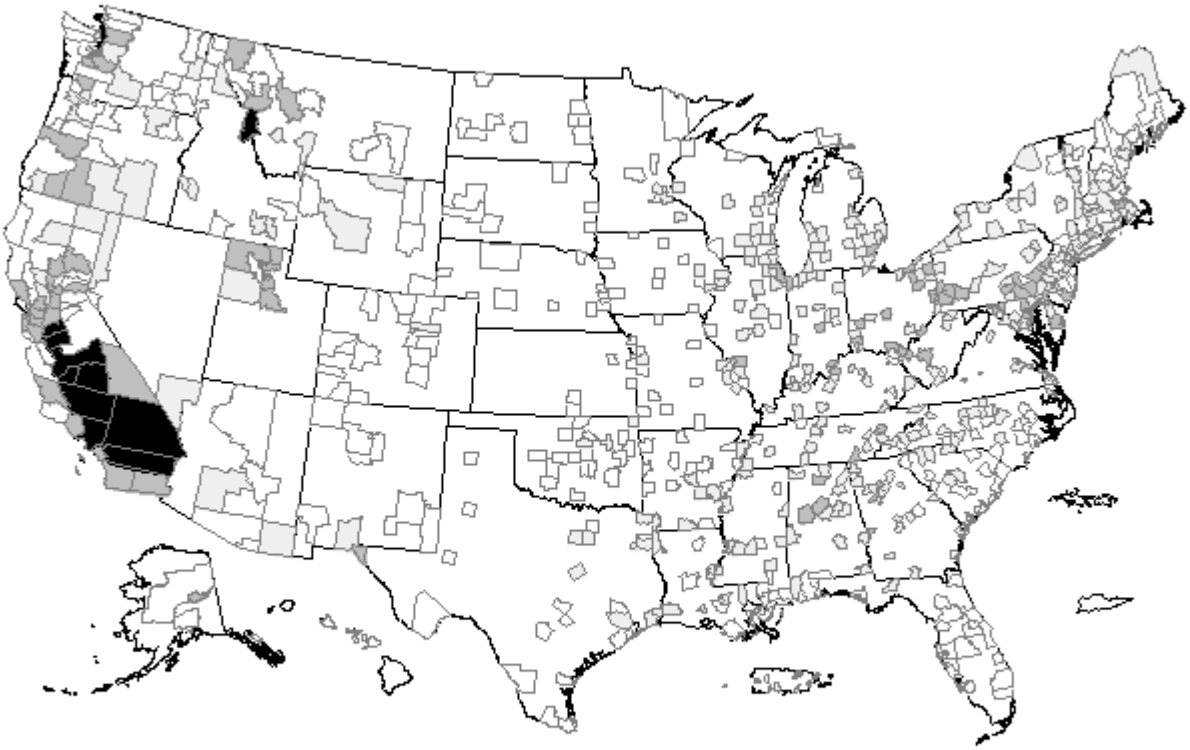
County Max Annual Mean $PM_{2.5}$ Concentrations (ug/m3), 2000-2002



Concentration ug/m3
0 < x <= 12
12 < x <= 15
15 < x <= 18
x > 18

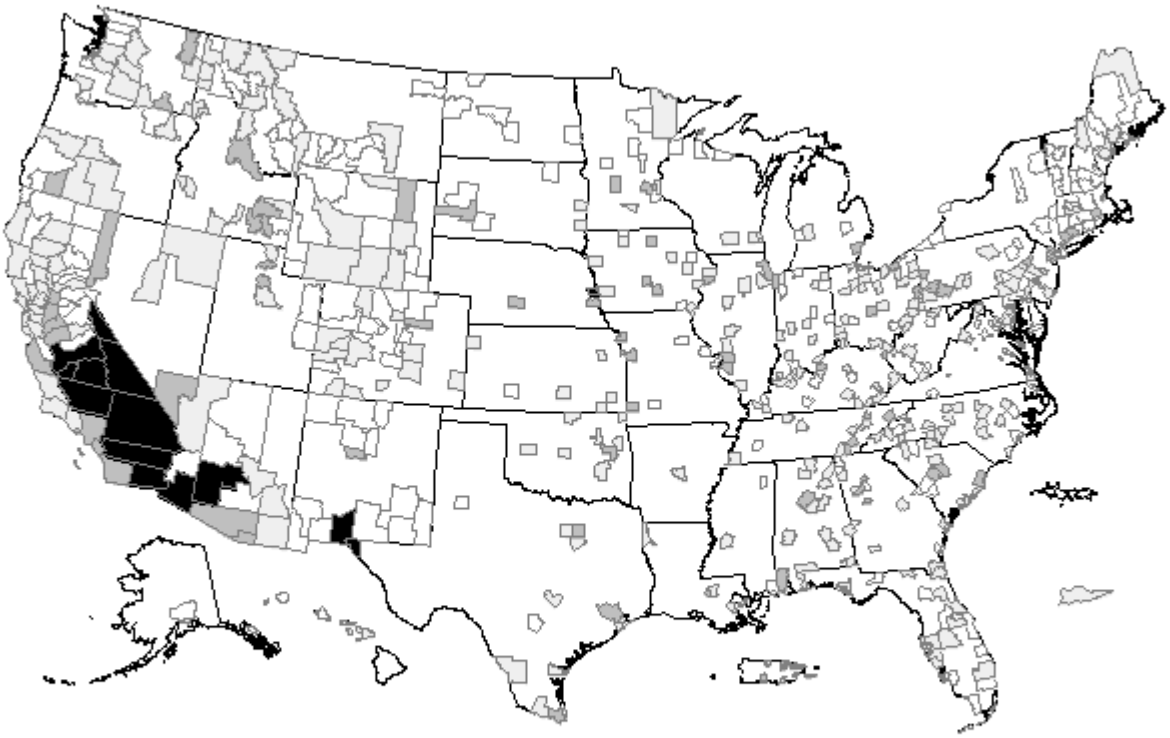
693 counties

County Max 98th Percentile 24-hour $PM_{2.5}$ Concentrations (ug/m3), 2000-2002



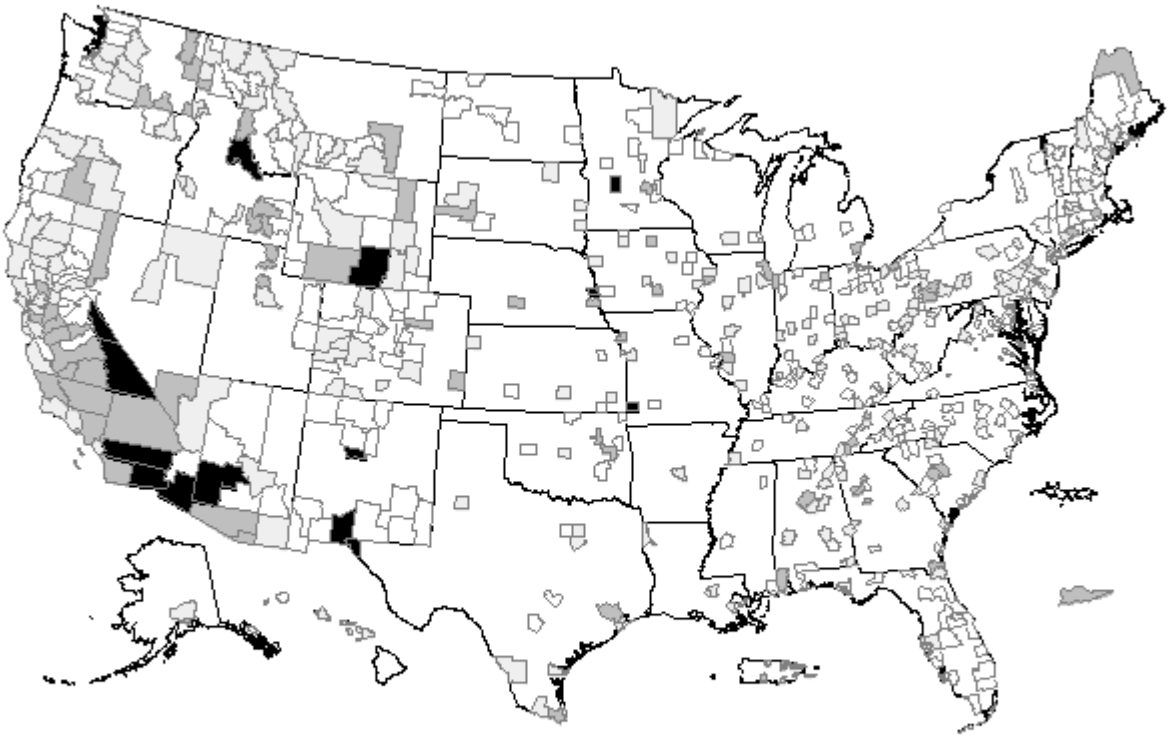
693 counties

County Max Annual Mean PM_{10} Concentrations (ug/m3), 2000-2002



574 counties

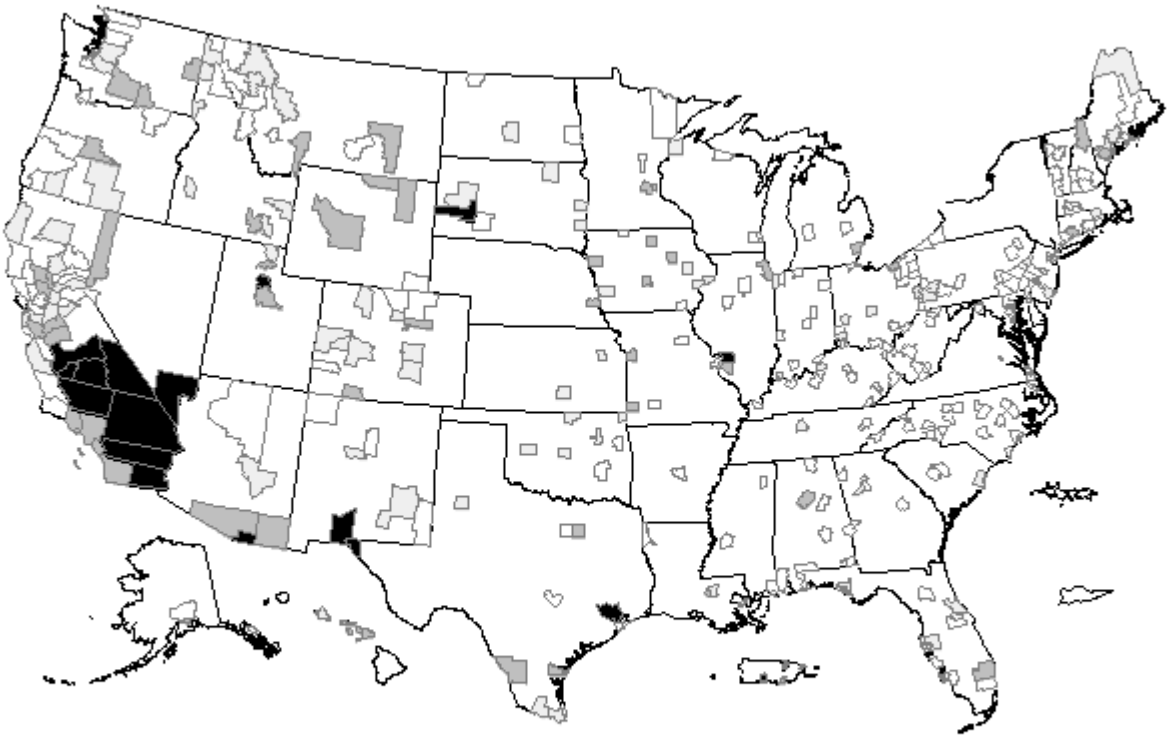
County Max 98th Percentile 24-hour PM_{10} Concentrations (ug/m3), 2000-2002



Concentration ug/m3 0 < x <= 50 | 50 < x <= 75 | 75 < x <= 150 | x > 150

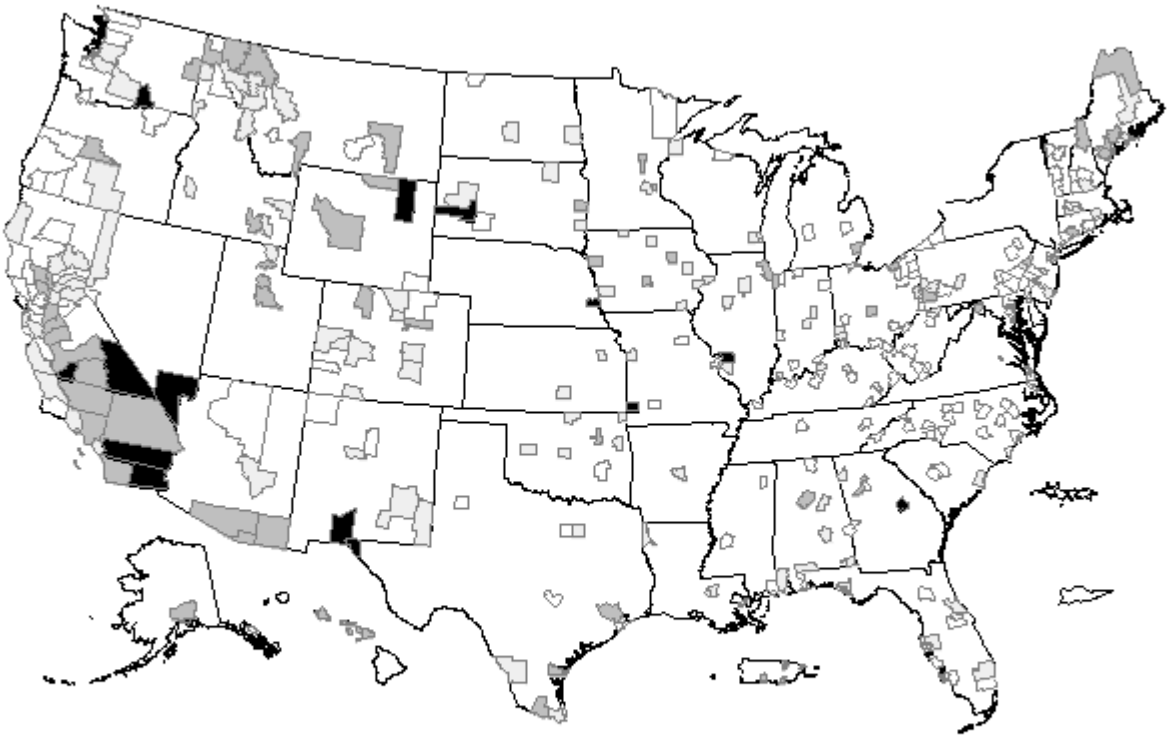
574 counties

County Max Estimated Annual Mean $PM_{10-2.5}$ Concentrations (ug/m3), 2000-2002



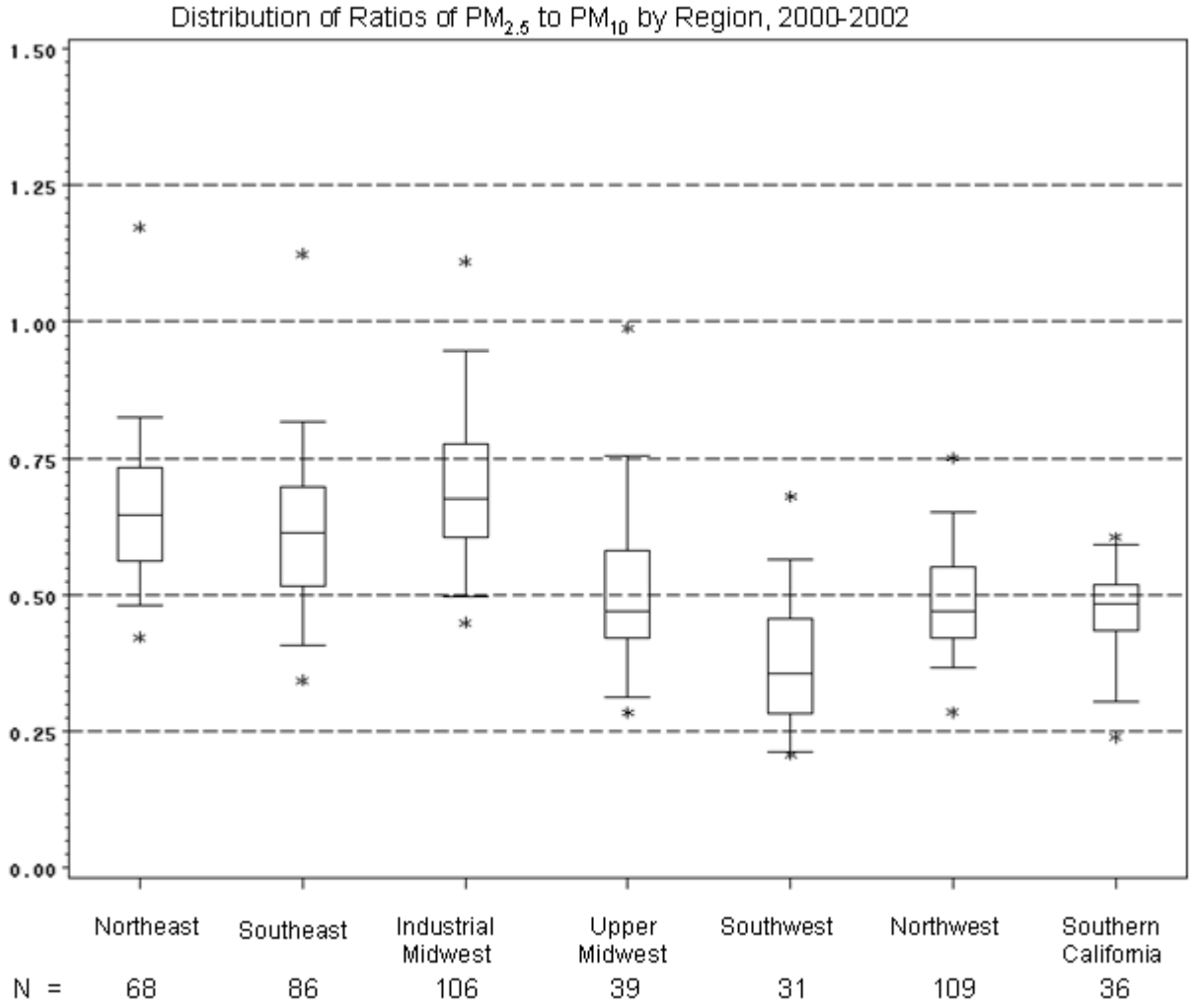
351 counties

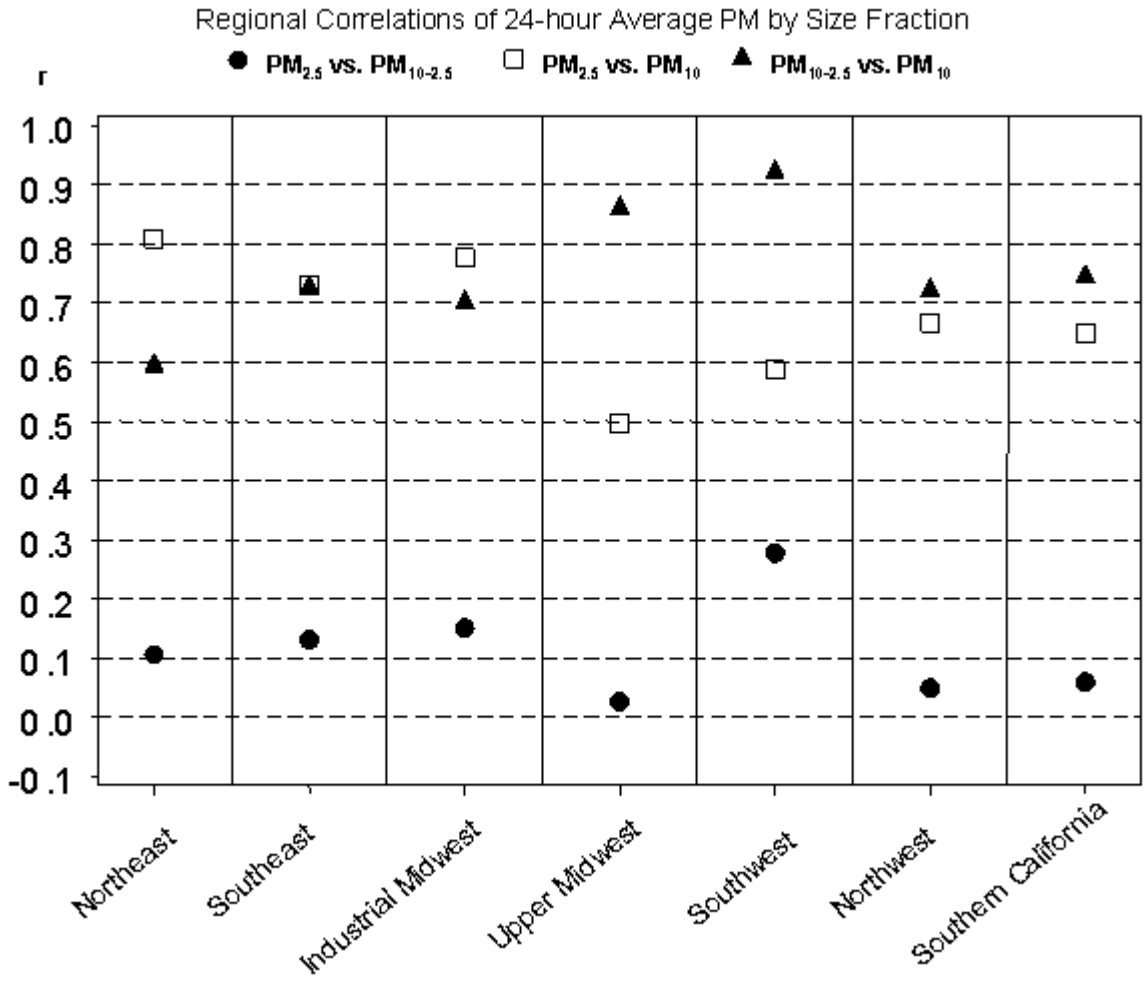
County Max Estimated 98th Percentile 24-hour $PM_{10-2.5}$ Concentrations (ug/m3), 2000-2002

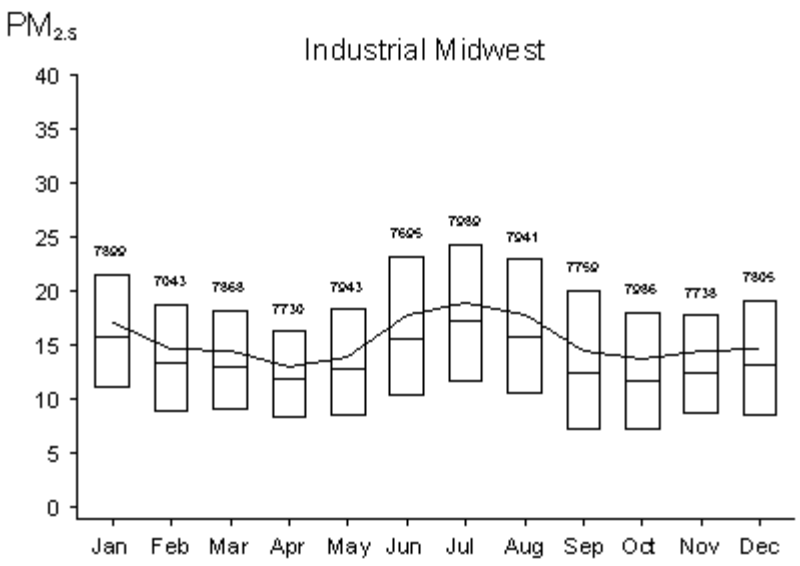


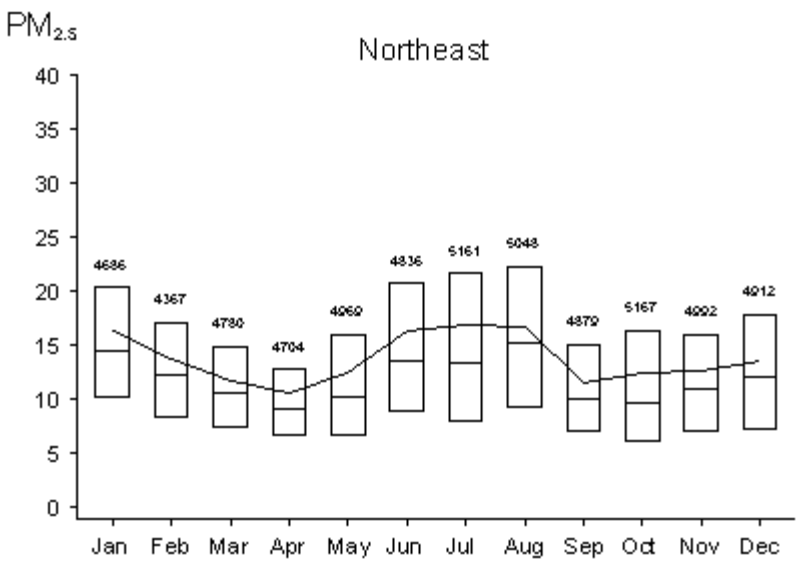
Concentration ug/m3 0 < x <= 25 | 25 < x <= 45 | 45 < x <= 75 | x > 75

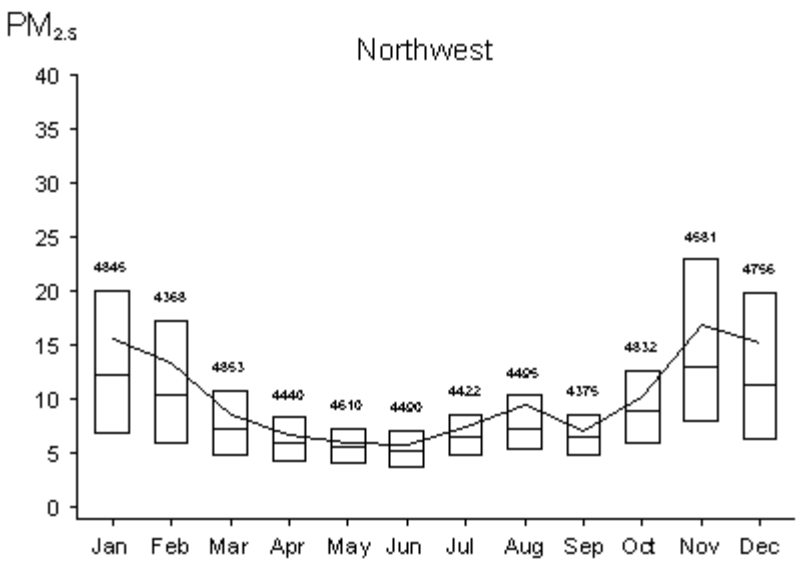
351 counties

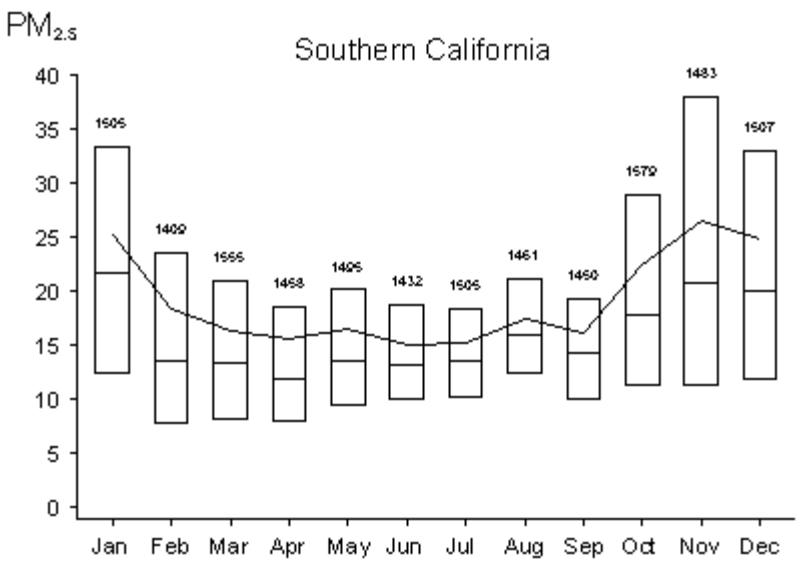

Distribution of Ratios of PM2.5 to PM10 by Region, 2000-2002
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Northeast
Southeast
Industrial Midwest
Upper Midwest
Southwest
Northwest
Southern California
N =
68
86
106
39
31
109
36

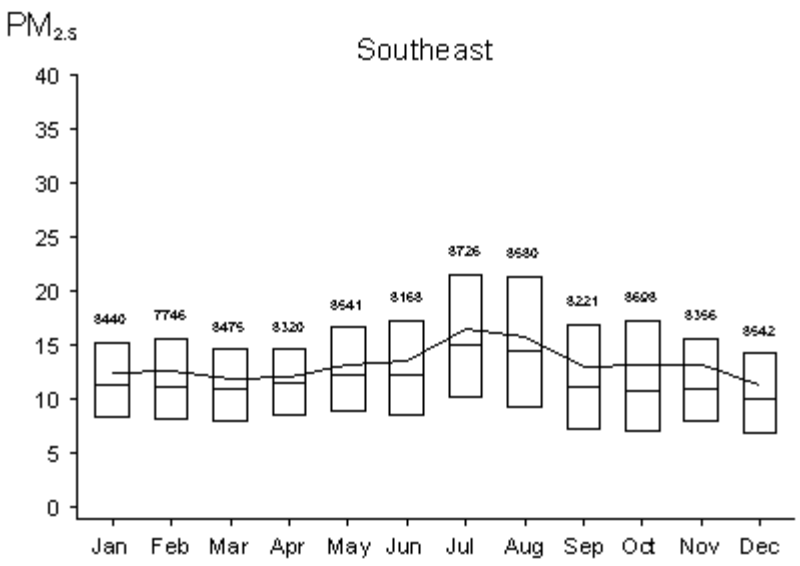

Regional Correlations of 24-hour Average PM by Size Fraction
r
PM2.5 vs. PM10-2.5
PM2.5 vs. PM10
PM10-2.5 vs. PM10
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
Northeast
Southeast
Industrial Midwest
Upper Midwest
Southwest
Northwest
Southern California

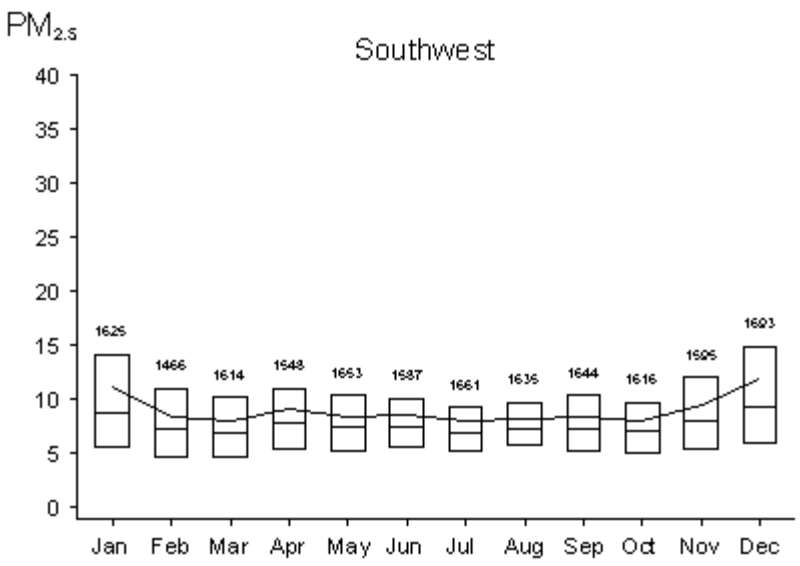

PM2.5
Industrial Midwest
40
35
30
25
20
15
10
5
0
7809
7043
7868
7730
7943
7695
7989
7941
7752
7286
7738
7805
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

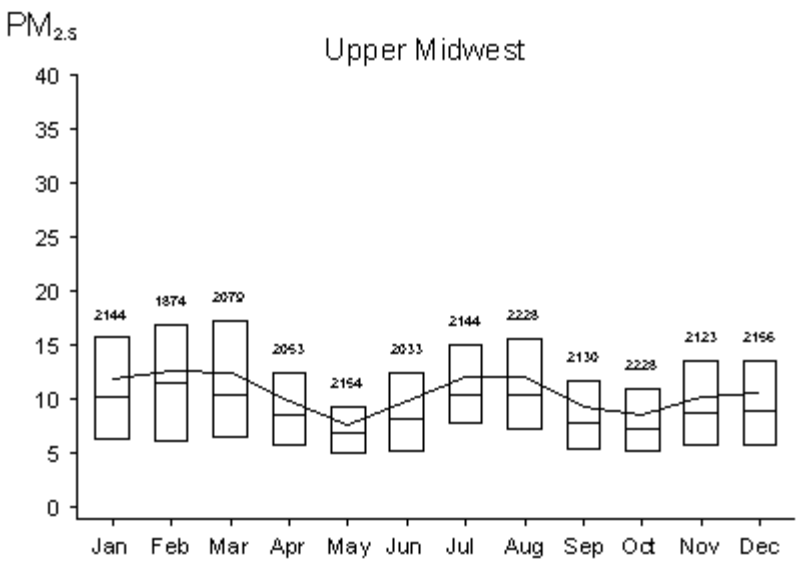

PM2.5
Northeast
40
35
30
25
20
15
10
5
0
4686
4367
4730
4704
4960
4836
5161
5048
4879
5167
4992
4912
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

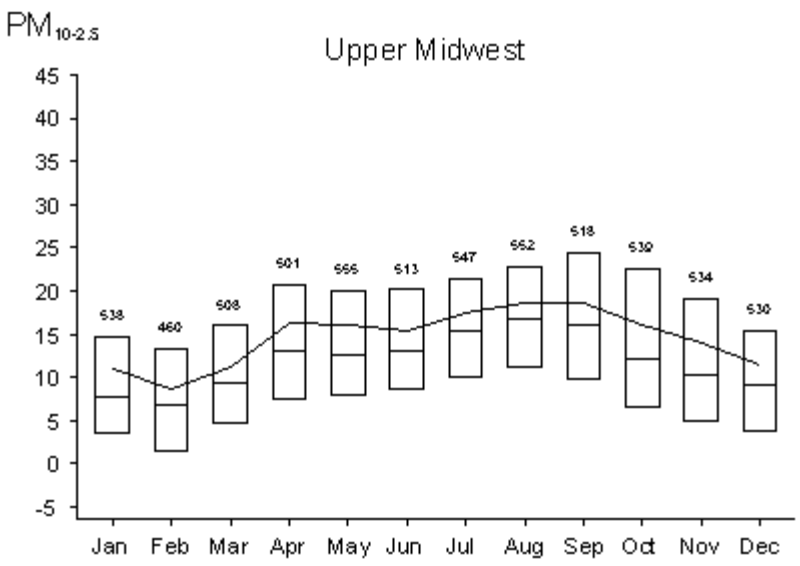

PM2.5
Northwest
40
35
30
25
20
15
10
5
0
4846
4368
4853
4440
4610
4490
4422
4496
4375
4832
4681
4756
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

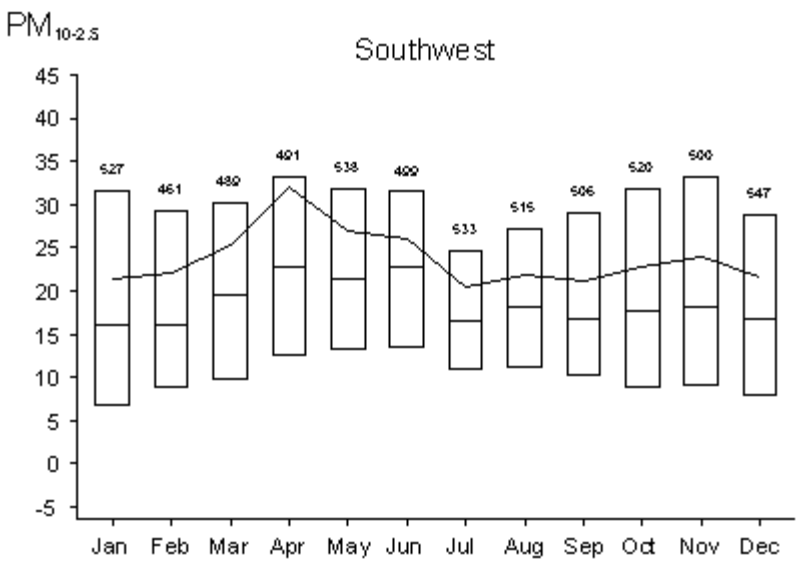

PM2.5
Southern California
40
35
30
25
20
15
10
5
0
1505
1402
1456
1458
1405
1432
1506
1461
1460
1579
1483
1507
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

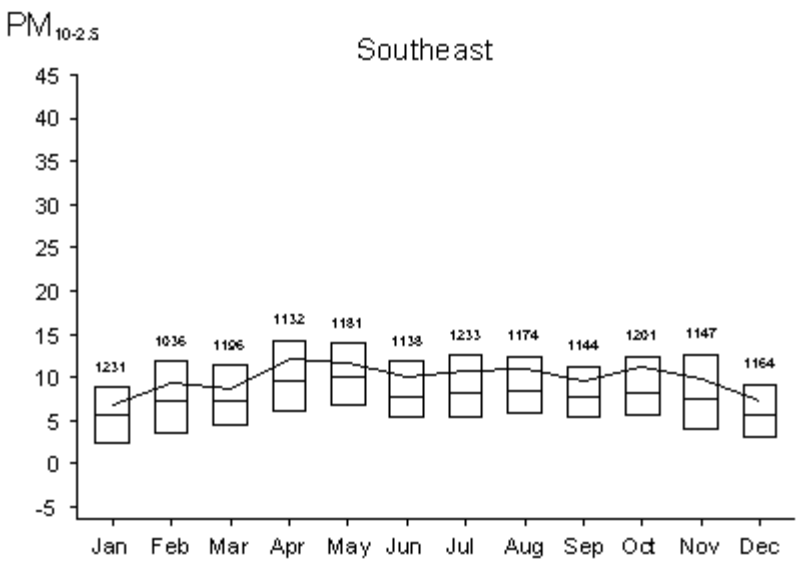

PM2.5
Southeast
40
35
30
25
20
15
10
5
0
8440
7746
8475
8320
8541
8168
8726
8680
8221
8608
8356
8542
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

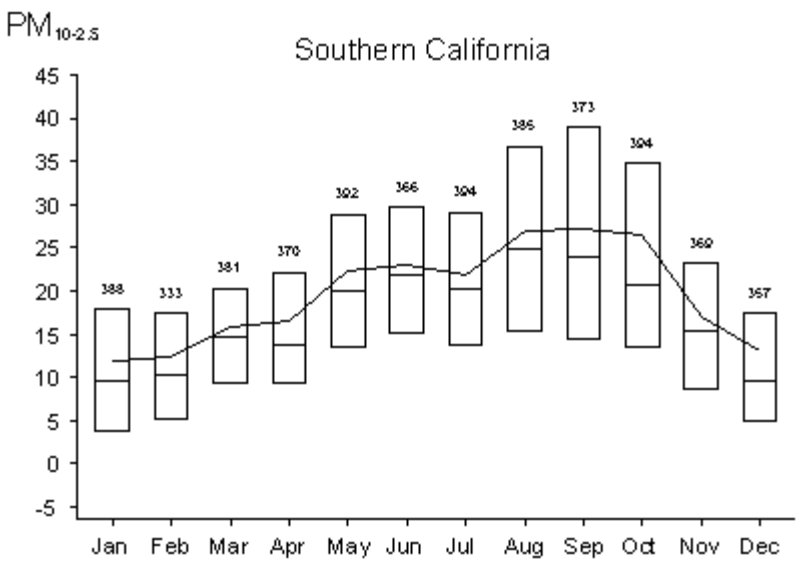

PM2.5
Southwest
40
35
30
25
20
15
10
5
0
1625
1466
1614
1648
1653
1587
1661
1635
1644
1616
1595
1693
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

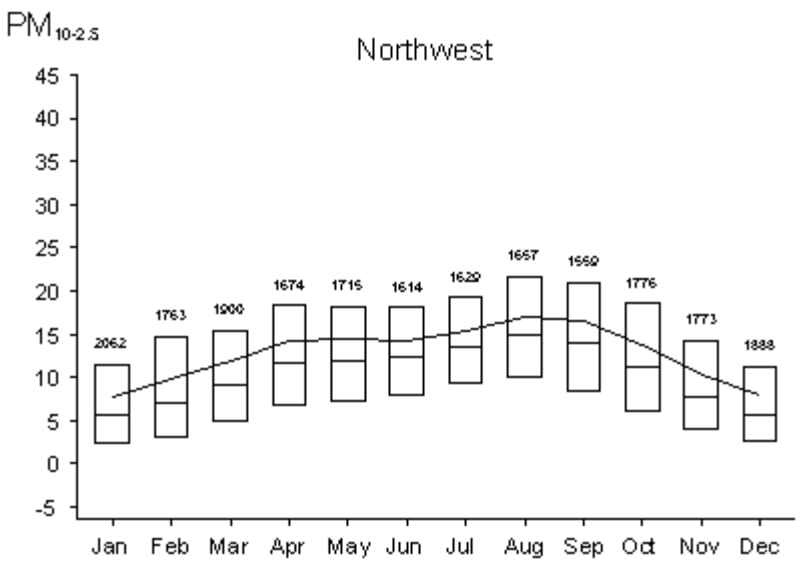

PM2.5
Upper Midwest
40
35
30
25
20
15
10
5
0
2144
1874
2079
2053
2154
2033
2144
2228
2130
2228
2123
2156
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

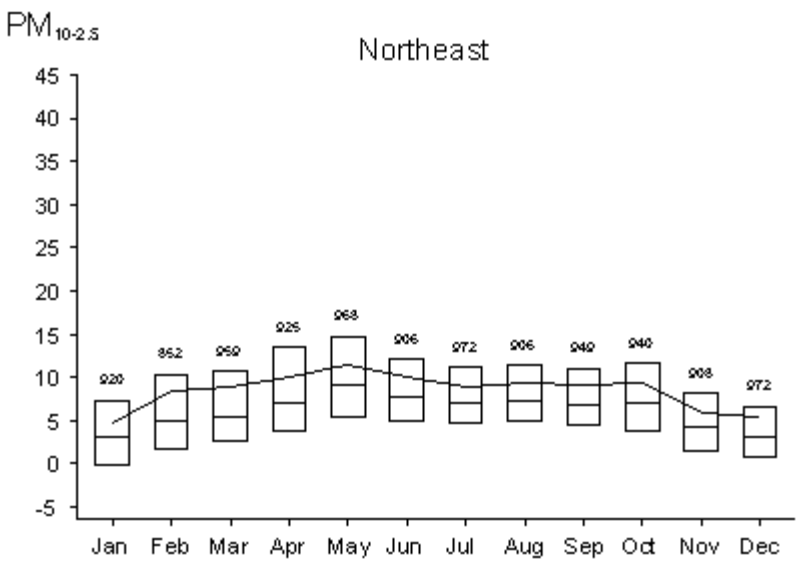

PM10-2.5
Upper Midwest
45
40
35
30
25
20
15
10
5
0
-5
538
460
508
601
555
513
547
562
618
539
534
530
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

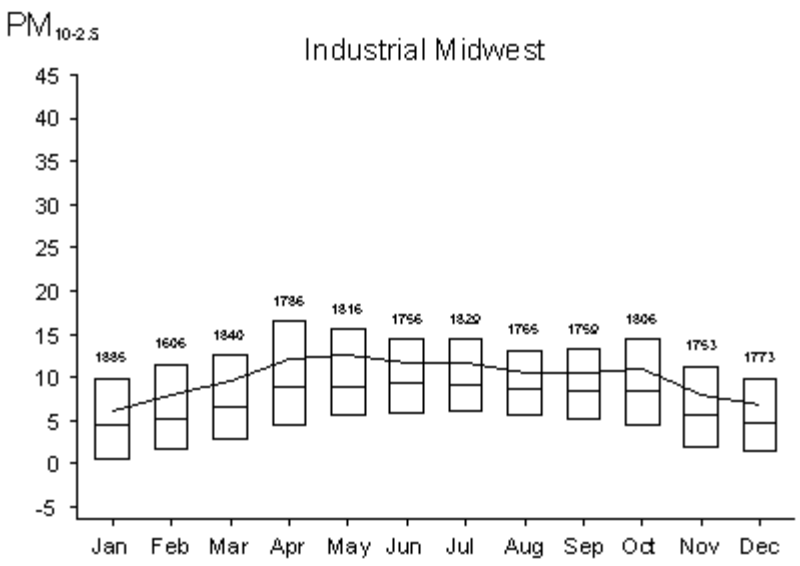

PM10-2.5
Southwest
45
40
35
30
25
20
15
10
5
0
-5
527
461
482
491
538
499
533
515
506
520
500
547
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec


PM10-2.5
Southeast
45
40
35
30
25
20
15
10
5
0
-5
1231
1036
1196
1132
1181
1138
1233
1174
1144
1201
1147
1164
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

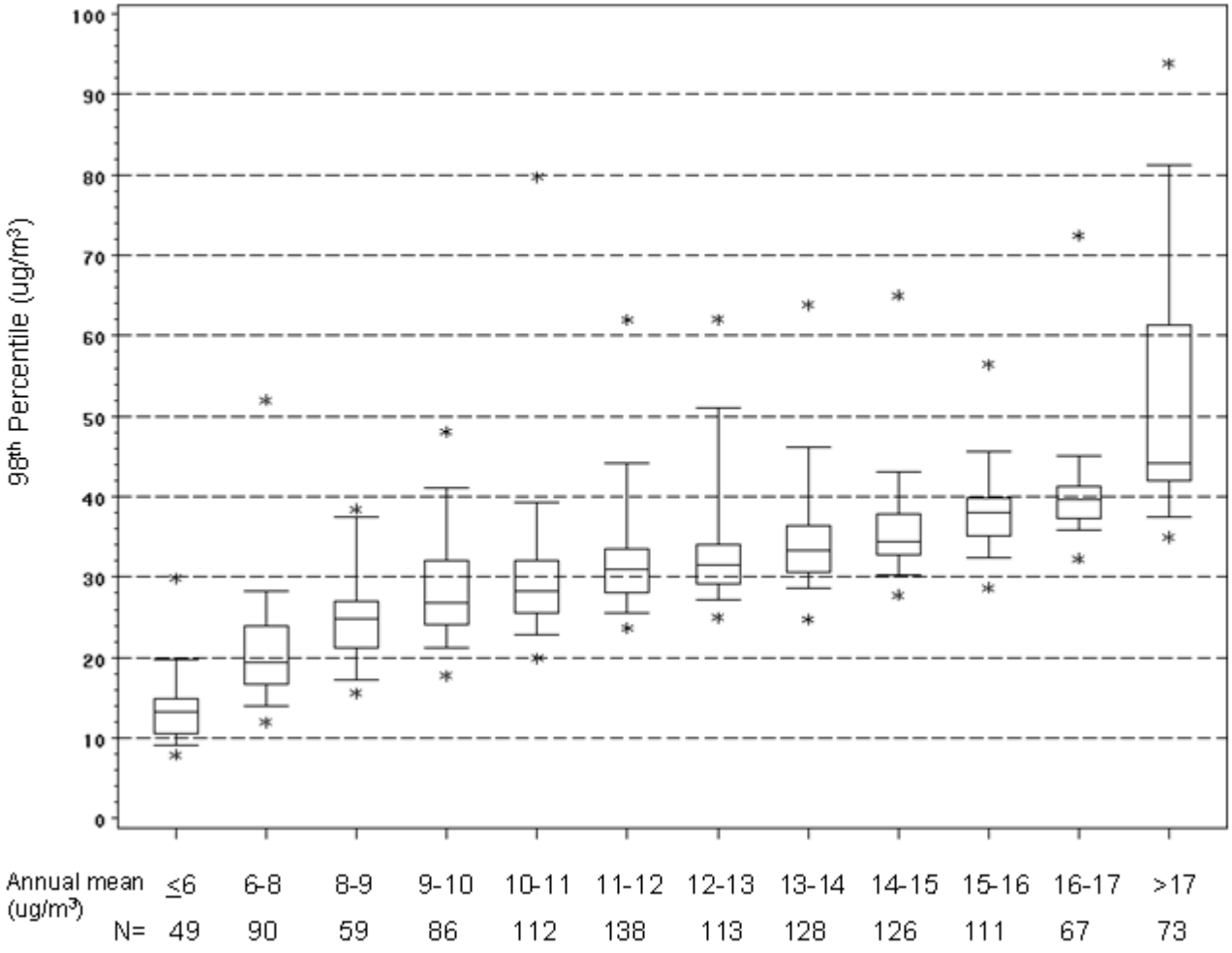

PM10-2.5
Southern California
45
40
35
30
25
20
15
10
5
0
-5
388
333
381
370
392
396
394
386
373
394
369
357
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec


PM10-2.5
Northwest
45
40
35
30
25
20
15
10
5
0
-5
2062
1763
1900
1674
1715
1614
1629
1657
1569
1776
1773
1888
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

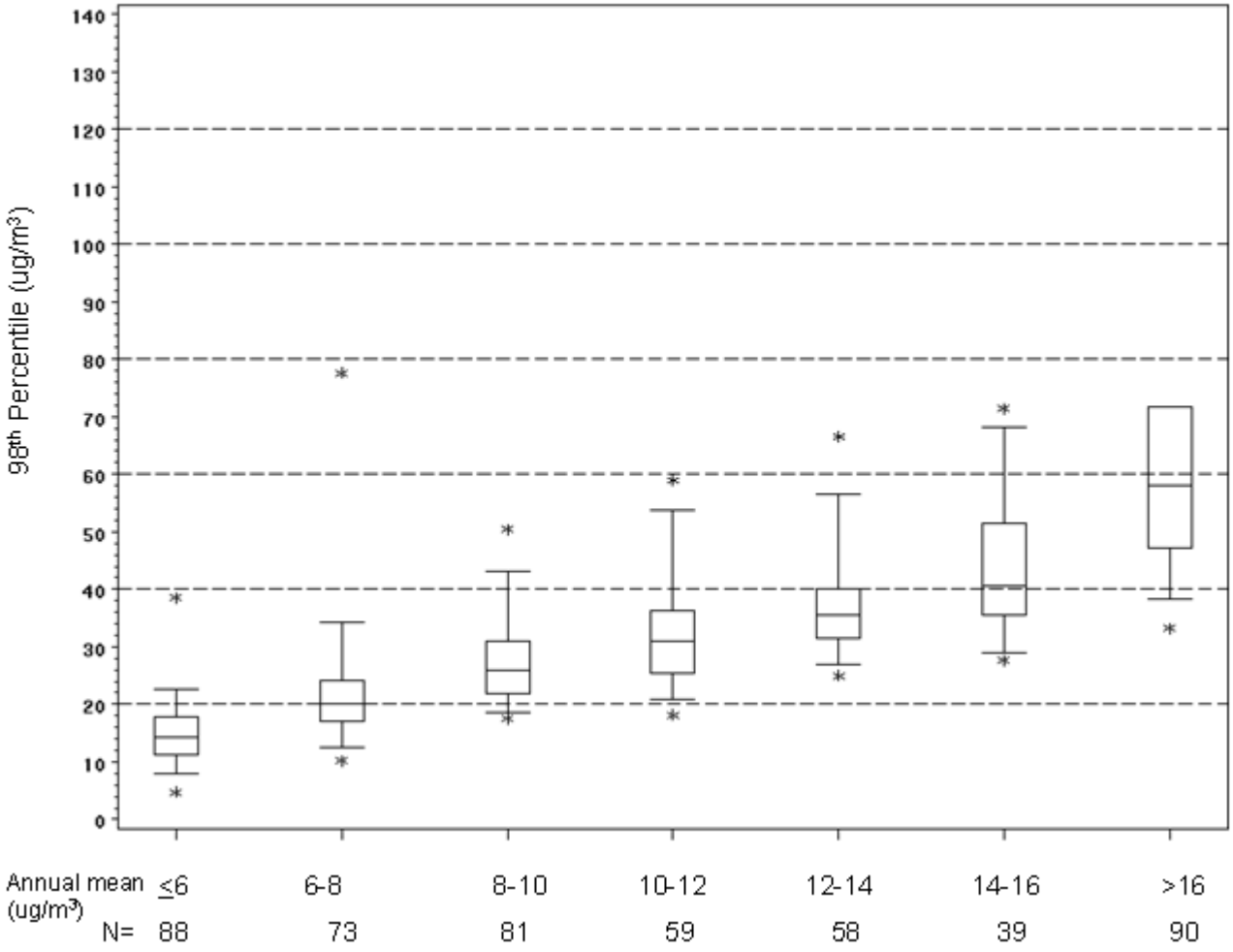

PM10-2.5
Northeast
45
40
35
30
25
20
15
10
5
0
-5
920
862
950
925
968
906
972
906
940
940
908
972
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec


PM10-2.5
Industrial Midwest
45
40
35
30
25
20
15
10
5
0
-5
1585
1606
1640
1786
1816
1756
1829
1766
1759
1806
1753
1773
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

Distribution of 98th Percentile $PM_{2.5}$ Concentrations by Annual Mean Interval, 2000-2002



Distribution of 98th Percentile $PM_{10-2.5}$ Concentrations by Annual Mean Interval, 2000-2002



ATTACHMENT B

Processing Details for AQS Hourly Sample Duration Figures

This attachment describes the data (2000-2002) and processing procedures used to generate the following AQS based, hourly sample duration figures:

- Figure 2-22. Hourly average $PM_{2.5}$ distributions (upper panel) and seasonal average $PM_{2.5}$ concentrations (lower panel) at a Cleveland, OH monitoring site, 2000-2002.
- Figure 2-23. Hourly average $PM_{10-2.5}$ distributions at a Cleveland, OH monitoring site, 2000-2002.

The graphics are included in the attached file, 'Attach-B-graphics'.

Data Description

Hourly (AQS duration='1') PM_{10} and $PM_{2.5}$ data for calendar years 2000-2002 were extracted from AQS on May 28, 2003. Two queries were made for PM_{10}: one for parameter 81102 [PM_{10}, standard temperature and pressure conditions (STP)] and one for parameter 85101 [PM_{10}, local temperature and pressure conditions (LTP)]. Only one query was necessary for $PM_{2.5}$, for parameter 88101 ($PM_{2.5}$ LTP). Hourly meteorological data were obtained from the National Weather Service (NWS). Meteorological data, specifically temperature and pressure, were necessary to convert PM_{10} data reported at STP to LTP. To convert the PM_{10} STP data to an LTP basis, the following formula was used:

$$PM_{10}\text{-LTP} = PM_{10}\text{-STP} * [298 / \text{temperature (degrees C)}] * [\text{pressure (mm Hg)} / 760].$$

Hourly $PM_{10-2.5}$ records were created by subtracting the $PM_{2.5}$ from the LTP (reported or converted) PM_{10} values. Since data are reported to AQS in Standard Time, hours were adjusted to reflect daylight savings time to be consistent with human activity. The seasons were derived in the following manner:

- December, January, February (quarter 1) = Winter
- March, April, May (quarter 2) = Spring
- June, July, August (quarter 3) = Summer
- September, October, November (quarter 4) = Fall

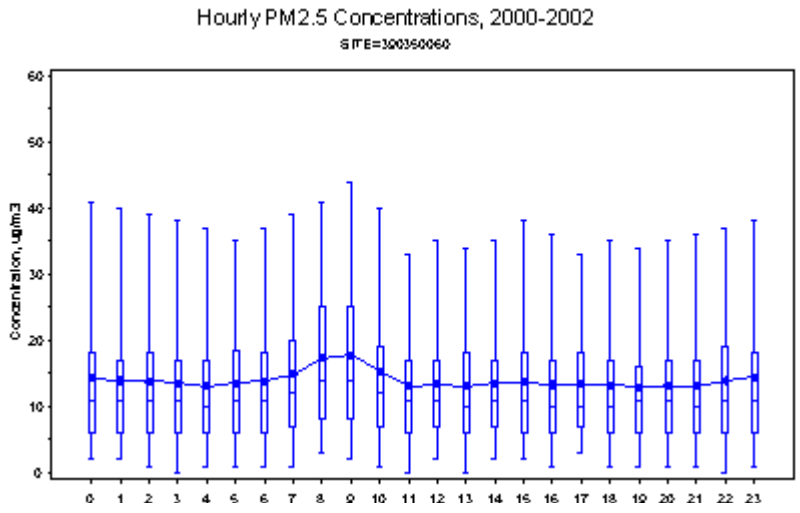
Outputs Description

Two types of plots are shown in the analyses figures, boxplots and a diurnal line plots. In the boxplots, the boxes represent the interquartile range (25^{th} to 75^{th} percentiles) of each hourly distribution, the line inside each box shows the median of the distribution, the dot inside each box represents the distribution mean (the distribution means are connected by lines), and the box whiskers represent the 5^{th} and 95^{th} percentiles. The diurnal line graph plots hourly averages by season. The hourly averages (by season) are joined by lines; plotting symbols are not shown. Although the outputs were produced for multiple sites, only the graphics for the Cleveland, OH

site (AQS ID = 390350060) were used in the analyses memorandum.

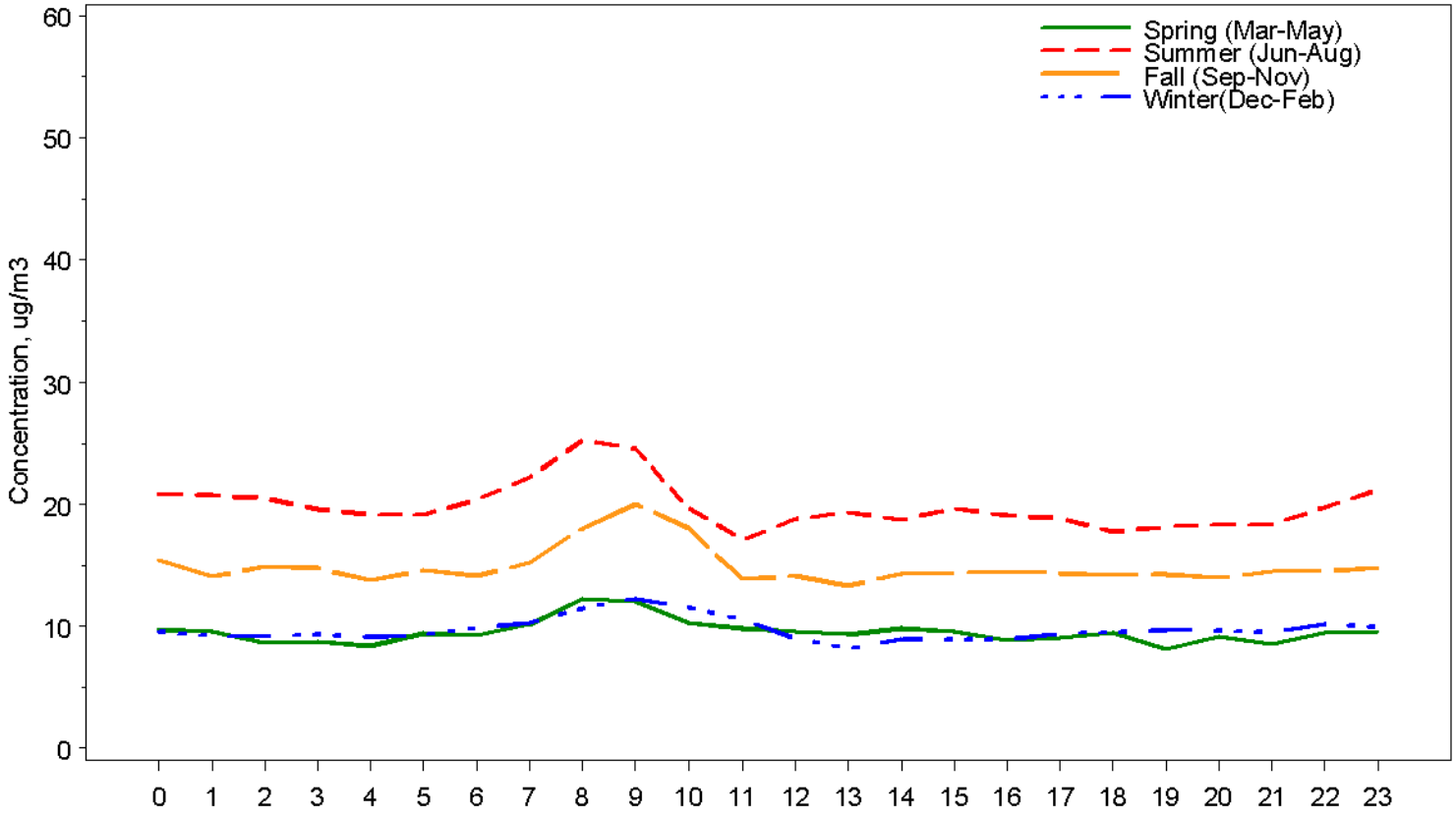
Methodology

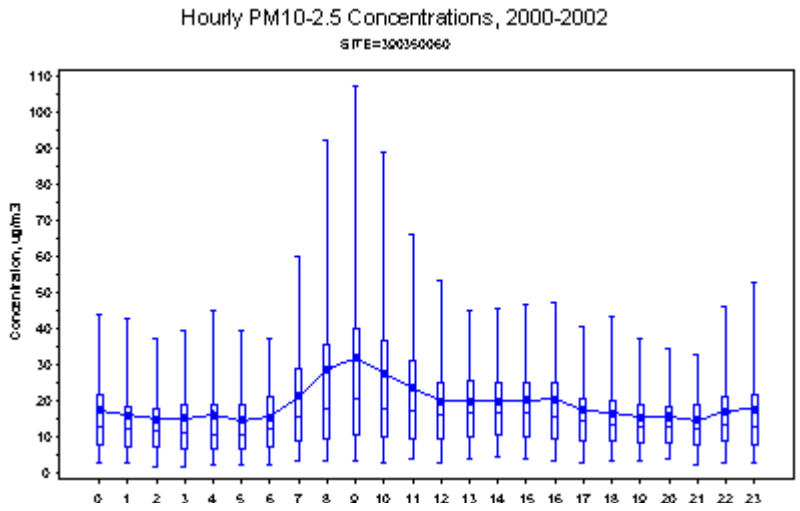
The statistical software package SAS was used to process the data and create the associated graphics. The SAS procedure PROC UNIVARIATE was used to calculate the summary statistics and the procedure PROC GPLOT was used to procedure the graphics. The code used to generate the plots are contained in Attach-B-SAScode.


Hourly PM2.5 Concentrations, 2000-2002
SITE=100350060
Concentration, ug/m3
60
50
40
30
20
10
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Average Hourly PM2.5 Concentrations, 2000-2002

MSA_NAME=Cleveland-Lorain-Elyria,OH




Hourly PM10-2.5 Concentrations, 2000-2002
SITE=300350060
Concentration, ug/m3
0 10 20 30 40 50 60 70 80 90 100 110
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

ATTACHMENT C

Processing Details for IMPROVE and ESpN Figures

This attachment describes the data and processing procedures used to generate the following ESpN and IMPROVE figures:

- Figure 2-8. Average annual mean $PM_{2.5}$ concentration trend at IMPROVE sites, 1999-2001.
- Figure 2-13. Average annual mean $PM_{10-2.5}$ concentration trend at IMPROVE sites, 1999-2001.
- Figure 2-14. Annual average composition of PM2.5 by Region. Upper panel rural sites, lower panel urban sites.
- Figure 2-15. Average annual mean sulfate, total carbon, and crustal material concentration trend at IMPROVE sites, 1999-2001.

The graphics are included in the attached file, 'Attach-C-Graphics'.

Data Description

Two types of data were used in these analyses: Rural particle data were derived from daily measurements from the Interagency Monitoring of PROtected Visual Environmental (IMPROVE) aerosol monitoring network. Urban aerosol particle data were derived from daily measurements from the EPA Speciation Monitoring Network (ESpN). Data from ESpN were extracted from AQS on July 3, 2003. Data from IMPROVE were obtained in two manners. Debbie Miller of the National Park Service provided 1992-2001 site-level annual summary data on July 10, 2003. Additional IMPROVE summary data for the time period September, 2001 through August, 2002 were downloaded from the IMPROVE website on July 20, 2003. All summary data used in these analyses are included in the attached file, 'Attach-C-Data'.

The IMPROVE trends analyses (Figures 2-8, 2-13 and 2-15) show 10-year trends for $PM_{10-2.5}$ mass; $PM_{2.5}$ mass; and the three $PM_{2.5}$ mass components: sulfate (SO_4), total carbon (TC, sum of organic carbon and elemental carbon), and crustal material. All five major components of $PM_{2.5}$ mass are reported in the IMPROVE / ESpN bar graphs of Figure 2-14. These components include: ammonium (NH_4), nitrate (NO_3), total carbonaceous mass (TCM), and crustal material. TCM is calculated as organic carbon mass (OCM) + elemental carbon (EC). OCM is estimated as measured and blank-corrected organic carbon (OC) multiplied by 1.40 (to convert to mass). Crustal material concentrations are estimated using the 'IMPROVE equation:

2.2[Aluminum]+2.49[Silicon]+1.63[Calcium]+2.42[Iron]+1.94[Titanium].

Outputs Description

The IMPROVE trends plots (Figures 2-8, 2-13 and 2-15) show the annual average concentrations for $PM_{10-2.5,}$ $PM_{2.5}$, sulfate, total carbon, and crustal material averaged across 9 eastern sites and 23 western sites. There is also a separate plot series for the Washington, DC site because it is not a rural site like the others. A list of sites by area classification is provided in the file 'Trends Region X Site.xls' (which is located in 'Attach-C-Data.zip'.)

The bar graphs contrast rural (top panel, data from IMPROVE) versus urban (bottom panel, data from ESpN) average concentrations of $PM_{2.5}$ constituents by geographic region. For this particular analysis, the following geographic regions were used (in lieu of HEI region definitions): South East, Mid West, East Coast/North East, East Texas/South, Far North East, North Plains, California, Desert West, and North West. See spreadsheet 'Bar Charts Region X Site.xls' (located in 'Attach-C-data.zip') for the regional site assignments.

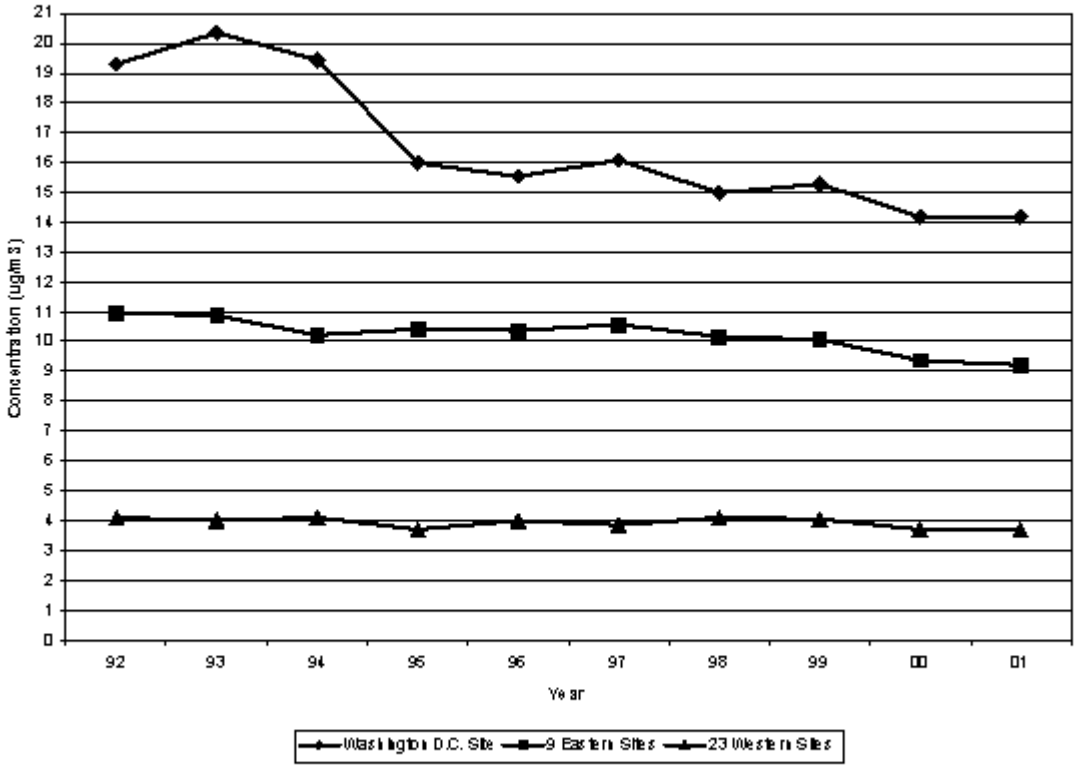
Methodology

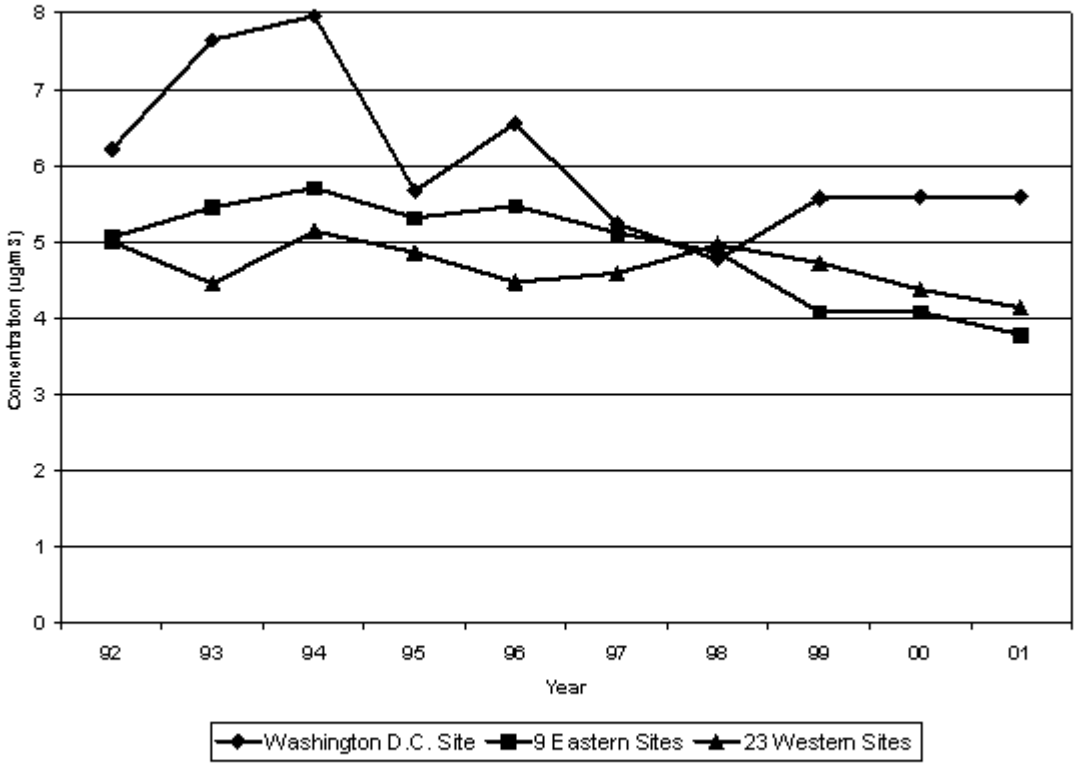
For the trends analyses, sites were required to have 8 of 10 valid years of data. Missing years were interpolated using surrounding years. The lines plotted are the averages across the trend sites in each region. The Washington plot is based on a single site. PROC UNIVARIATE in the SAS statistical software package was used to calculate the averages in each plot. The SAS code is included in the attached file, 'Attach-C-SAScode'. Plotting points for each of the three plots are presented in file, 'Trends Plotting Points.xls' which is located in 'Attach-C-Data.zip'.

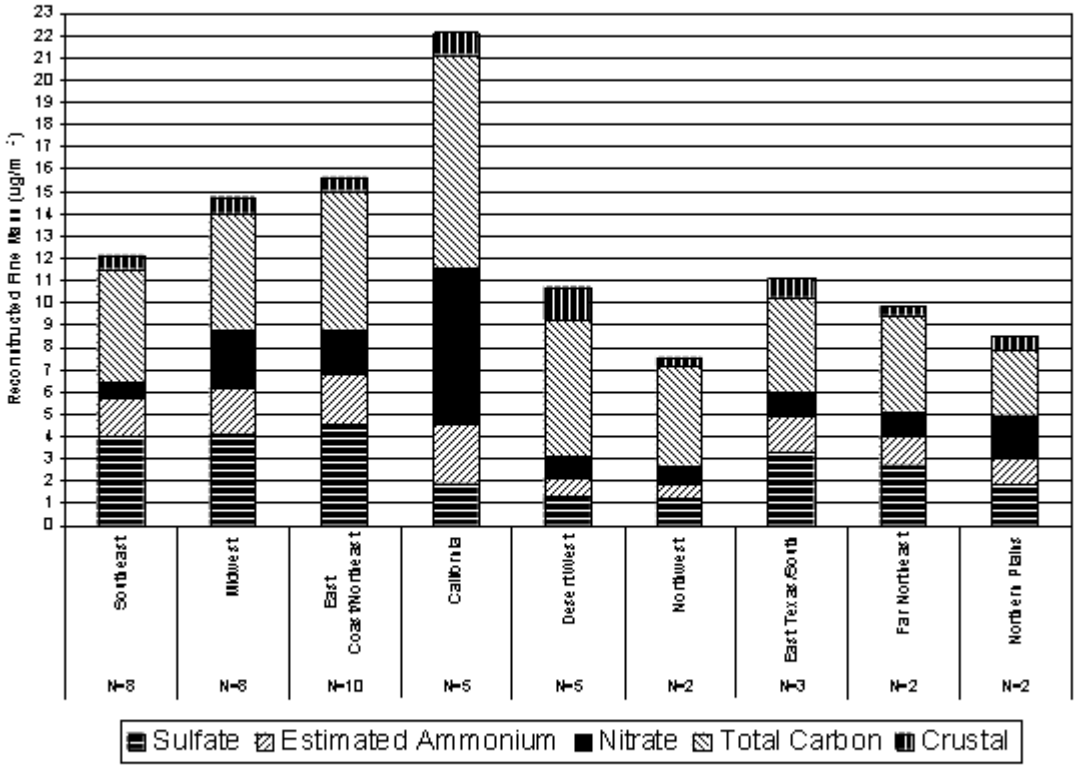
For the bar graphs, only data for 'complete' sites were used. For both the urban and rural data, 'complete' data consisted of having 50% or more observations (of the major chemical components of $PM_{2.5}$ mass: sulfate, nitrate, organic carbon mass, elemental carbon, aluminum, calcium, titanium, iron, and silicon) per quarter for the year analyzed

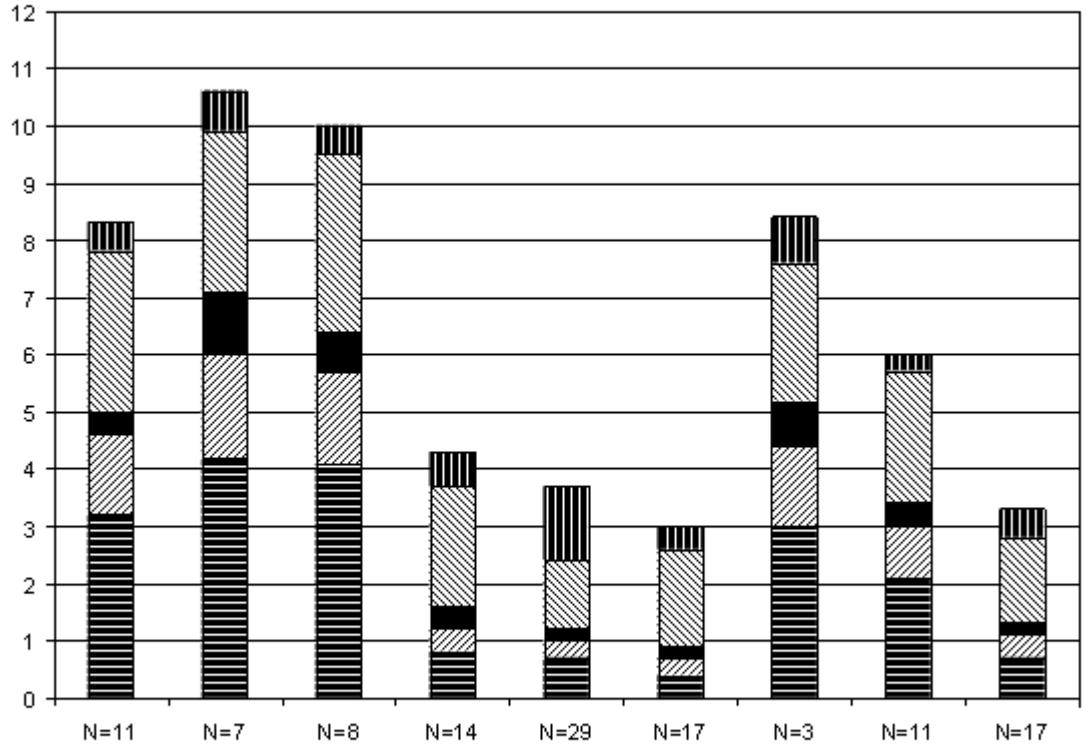
Use of Outputs

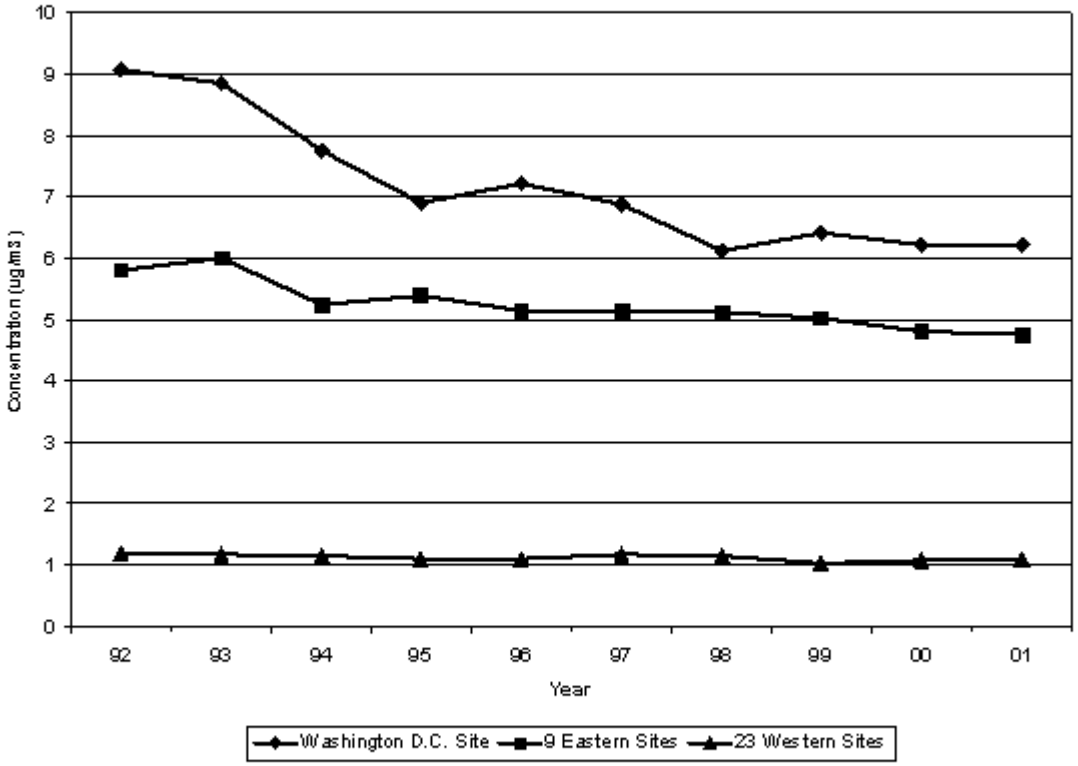
The data as presented in the bar charts should be used only to gauge relative levels of urban versus rural concentrations of the chemical species in each of the regions. Because urban and rural sites were not specifically matched (in terms of separation distance, representative upwind locations, etc.), no inferences should be drawn about 'urban increments' based on these data and graphics. The two bar charts simply illustrate how $PM_{2.5}$ constituents vary for a group of urban and rural sites that are roughly located in the same geographic region. For further details on 'urban increment' analyses, refer to: V. Rao, N. Frank, A. Rush, F. Dimmick, "Chemical Speciation of PM2.5 in Urban and Rural Areas", In the *Proceedings of the Air & Waste Management Association Symposium on Air Quality Measurement Methods and Technology*, San Francisco, November 13-15, 2002.

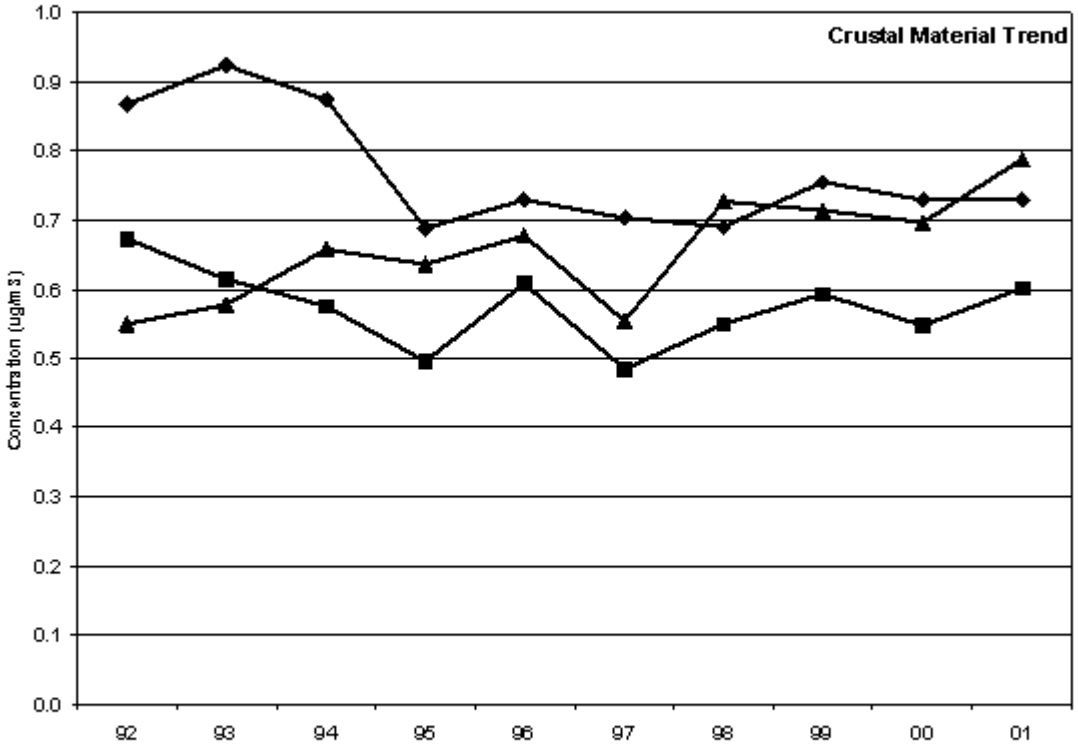

Concentration (ug/m3)
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
92
93
94
95
96
97
98
99
00
01
Year
Washington D.C. Site
9 Eastern Sites
23 Western Sites

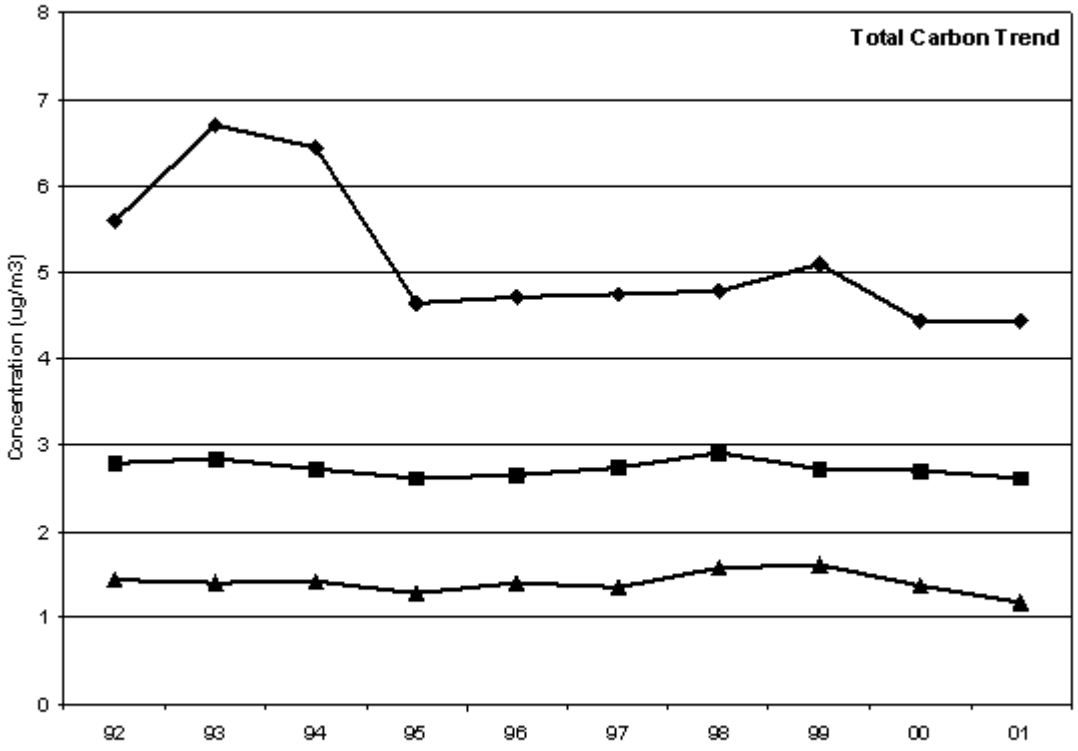

8
7
6
5
4
3
2
1
0
Concentration (ugm3)
92
93
94
95
96
97
98
99
00
01
Year
Washington D.C. Site
9 Eastern Sites
23 Western Sites


Reconstructed Fine Mass (ug/m^3)
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
Southeast
Midwest
East Coast/Northeast
California
Desert/Midwest
Northwest
East Texas/South
Far Northeast
Northern Plains
N=8
N=8
N=10
N=5
N=5
N=2
N=3
N=2
N=2
Sulfate
Estimated Ammonium
Nitrate
Total Carbon
Crustal


12
11
10
9
8
7
6
5
4
3
2
1
0
N=11
N=7
N=8
N=14
N=29
N=17
N=3
N=11
N=17


10
9
8
7
6
5
4
3
2
1
0
Concentration (ug/m3)
92
93
94
95
96
97
98
99
00
01
Year
Washington D.C. Site
9 Eastern Sites
23 Western Sites


Crustal Material Trend
Concentration (ug/m3)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
92
93
94
95
96
97
98
99
00
01


Total Carbon Trend
Concentration (ug/m3)
0
1
2
3
4
5
6
7
8
92
93
94
95
96
97
98
99
00
01