Attachment A

AQS-Based, 24-Hour Duration PM Analyses

General / Background:

This attachment describes the analyses of 24-hour duration $PM_{2.5}$ and PM_{10} data obtained from AQS. It also documents the analyses of 24-hour duration $PM_{10-2.5}$ estimates which were largely derived from the aforementioned AQS datasets; a limited amount of directly reported $PM_{10-2.5}$ data (via dichotomous samplers) were also obtained directly from AQS.

Construction of $PM_{2.5}$ database

The database utilized for most 24-hour $PM_{2.5}$ FRM PM Staff Paper (SP) analyses (referenced subsequently as the 'regular' database) is a hybrid of the one used to construct 2001-2003 production design values (PDV's) and, hence, used in the PM25 designations process. Although the raw data are exactly the same, there are several core differences in the definition and determination of 'complete' sites. For the SP analyses, any site with 11 or more observations in all 12 quarters (2001-2003) was considered 'complete' and usable for general characterization. For PDV processing, 11 or more observations per quarter (all 12 quarters) was initially only sufficient (i.e., deemed a site 'complete') to prove nonattainment of the annual standard. To initially be deemed complete in order to show that the annual standard was being met, a site needed at least 75% data capture in all 12 quarters; the 75% cut-point was based on the required sampling frequency. Additionally for the PDV processing, sites that initially did not meet the required completeness goals (11+ samples or 75%+ capture) but were close were then subjected to several conservative data 'substitution' routines to see if there was a good likelihood that they would have shown attainment or nonattainment of the standard had they actually met the completeness goals. These substitution routines included the substitution (for evaluation purposes only, not for actual modification of their PDV's) of low values for missing data to show nonattainment, and high values for missing data to show attainment. Sites that passed one or more of these tests were then deemed complete and their PDV considered valid. For SP analyses, data substitution was not implemented. One additional difference between PDV and SP processing is the treatment of flagged data. Unless otherwise specified, all data including flagged event (exceptional or natural) data were used for general SP characterization analyses. For PDV processing, regionally-concurred event-flagged data were excluded from the official design values, although such data did count towards completeness requirements. It should be noted that in both PDV processing and the SP analyses, the 3-year average metrics (annual mean and 98th percentiles for both, and 99th percentile for SP analyses) are referred to as design values (DV's). Separately, the 3-year DV's are frequently referred to simply as 'annual means' or '98th percentiles'. To reiterate and elaborate, for general SP characterization analyses, the following bullets are applicable (*unless otherwise noted for specific analyses*):

- 24-hour duration data for the time period 2001 to 2003 were polled from AQS for parameter 88101 [$PM_{2.5}$, local temperature and pressure conditions (LC)] on July 6, 2004.
- Only FRM data were considered. The following AQS method codes are considered FRM: 116, 117, 118, 119, 120, 123, 142, 143, 144, and 145.
- Event-flagged data were not omitted. (Event-flagged data were subsequently omitted for Analysis 3.)
- Data were processed on a site basis. The monitor with first occurring parameter occurrence code (POC) was considered the 'primary' monitor. If an additional monitor (POC) at the site (i.e., a 'collocated' one) contained an FRM sample on a day in which one was not present at the primary POC, then those data were used for the site. Essentially, all POC's were merged but only one sample per day maximum utilized, precedence given to the lowest POC number.
- SAS code ('pmfinemacro part1.sas') was used to pull the raw data from AQS; weed out non-FRM measurements; merge collocated monitor data to a site basis; ascertain data capture statistics; and compute means, percentiles, and corresponding DV's.
- To be considered complete and hence, usable for SP analyses, a site needed at least 11 samples in each of the 12 quarters (irregardless of sampling frequency). 827 sites met the completeness goal.
- Unless otherwise noted, the SP $PM_{2.5}$ database was used to generate the $PM_{2.5}$ plots, tables, and related outputs.. [Occasionally, the $PM_{2.5}$ component of the $PM_{10-2.5}$ database was used in order to enhance the $PM_{2.5}$ versus $PM_{10-2.5}$ comparisons.]

Construction of $PM_{10-2.5}$ databases

Two 24-hour $PM_{10-2.5}$ databases (db's) were generated for SP-related analyses, a core database (termed the 'regular' db) and a somewhat larger database (called the 'extended' db). The regular db was utilized for almost all $PM_{10-2.5}$ analyses. The extended db was only used for estimating the number/percentage of counties that would potentially not meet alternative NAAQS levels. The extended db is defined later in this document when that specific analysis is described; the regular db is described here. In general, the regular db was constructed largely from collocated, same-day FRM/FEM PM_{10} and $PM_{2.5}$ measurement pairs utilizing a simplistic difference computation. This element of the processing was very similar to that implemented for the 2003 SP processing. However, this time an additional $PM_{10-2.5}$ component was included, that being direct measurements of the size cut emanating from dichotomous ('dichot') samplers. The following statements detail the $PM_{10-2.5}$ db construction:

- 24-hour duration data for the time period 2001 to 2003 were reterieved from AQS for the following parameters on August 24, 2004: parameter 88101 ($PM_{2.5}$, LC); parameter 81102 [PM_{10}, standard temperature and pressure conditions (STP)]; parameter 85101 (PM_{10}, LC); parameter 86101 ($PM_{10-2.5}$, STP); and parameter 81103 ($PM_{10-2.5}$, STP)
- Summary daily data (which includes hourly measurements aggregated within AQS to a 24-hour period) were extracted from AQS (also on August 24, 2004) for parameter 81102 and parameter 85101. AQS maintains the raw hourly data and also

aggregates the hourly information into summary daily records. A summary record is only deemed 'valid' if 75% or more of the hourly data (≥18) are present.

- For the difference method, only FRM/FEM PM_{10} and $PM_{2.5}$ data were utilized. All AQS PM_{10} data (except for a lone site in Alabama, ID '010970030') were assumed to be FRM/FEM. $PM_{2.5}$ data were determined to be FRM/FEM based on method code (as indicated above for $PM_{2.5}$ db development).
- For the difference method, no effort was made to account for differences in sampling instruments or protocols between the collocated 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.
- For the difference method, both the PM_{10} and $PM_{2.5}$ data used in the difference calculation were in units of µg/m3 at local conditions, thus the calculated $PM_{10-2.5}$ values also are in those units. 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 LC and not used in the analyses.
- For the difference method, $PM_{10-2.5}$ estimates were constructed from all site-day pairs of collocated PM_{10} and $PM_{2.5}$ measurements. For example, if, for a particular site day, there were two readings of PM_{10} ('1' and '2') and two readings of $PM_{2.5}$ ('a' & 'b', then four total $PM_{10-2.5}$ estimates were generated ('1a', '1b', '2a', and '2b).
- In situations where multiple site-day estimates of $PM_{10-2.5}$ existed (combination of difference method pair estimates and/or direct dichot measurements), they were averaged to obtain an average $PM_{10-2.5}$ measurement for the site-day. This average was considered the actual $PM_{10-2.5}$ estimate or 'sample' for that site-day (and counts as only one observation towards data completeness). Thus, data were essentially processed on a 'site' basis.
- Data for verified micro/middle-scale source-impacted $PM_{2.5}$ sites were eliminated from consideration into the potential $PM_{10-2.5}$ database; these sites were not considered to be appropriate candidates for future $PM_{10-2.5}$ network sites. The nine such sites are (AQS ID): '180890022', '180970066', '180970043', '170311016', '171190023', '170990007', '440070020', '481410053', and '291250001'.
- To be used in the SP analyses, a site needed 4, 8, or 12 consecutive quarters (2001-2003) of 11+ samples. This requirement is in contrast to the individual $PM_{2.5}$ and PM_{10} analyses which both required 'completeness' in all 12 quarters; the $PM_{10-2.5}$ criteria are more relaxed, in order to maximize the number of usable sites. Though nationally and regionally there are a sufficient number of 12-quarter complete PM25 sites and also a sufficient number of 12-quarter complete PM_{10} sites, there are <u>not</u> a sufficient number of *collocated* 12-quarter complete $PM_{2.5}$ and PM_{10} sites, Specifically, the $PM_{10-2.5}$ analyses utilized <u>the most recent</u> 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 2002, quarter 1. An additional increment of 4 consecutive quarters meets the 11 minimum sample threshold (2001, quarters 1-4), but would not have been used since a more recent band of data (shaded) were available. Although the utilized selection criteria do not guarantee a calendar year(s) of data, they do 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 subsequent analyses.

Year / Quarter =	2001, Q1	2001, Q2	2001, Q3	2001, Q4	2002, Q1	2002, Q2	2002, Q3	2002, Q4	2003, Q1	2003, Q2	2003, Q3	2003, Q4
N=	12	13	14	15	10	15	16	14	15	13	11	9

- 489 sites met completeness requirements: 137 with 4 complete quarters, 122 with 8 complete quarters, and 230 with all 12 complete quarters.
- ‘Annual’ means and percentiles (e.g., 98^{th}, 99^{th}) were computed from ‘annualized’ (4-quarter increment) statistics. For example, if a site had 8 complete quarters starting with 2001-Q3 and ending with 2003-Q2, then two ‘annual’ 98th percentiles were computed, one for 2001-Q3 through 2002-Q2 and the other for 2002-Q3 through 2003-Q2. Likewise, two ‘annual means’ were calculated (according to standard weighted mean processing protocol in which data are first averaged by quarter, and then the 4 quarterly means are averaged together to obtain an ‘annual mean’ figure). The 2 ‘annual’ numbers (2 means and 2 98^{th} percentiles) were then averaged to obtain the site’s DV-type metrics. Hence, the DV-type metrics might represent 4, 8, or 12 quarters of data. Separately, the (4, 8, or 12-quarter) DV-type metrics are frequently referred to simply as ‘annual means’ or ‘98^{th} percentiles’.
- Event-flagged data were not omitted. (Event-flagged data were subsequently omitted for Analysis 3.)
- SAS code was used to pull the raw 24-hour data from AQS (‘raw from aqs.sas’); extract the summary daily data from AQS (‘daily from aqs.sas’); process the AQS and NWS meteorological data needed to convert STP PM_{10} and $PM_{10-2.5}$ dichot data to LC (‘gettemppress0103.sas’); filter out non-FRM $PM_{2.5}$ data, create $PM_{10-2.5}$ difference records, and create an interim db of all site-day record (‘calc coarse 0103.sas’); average multiple site-day measurements, evaluate completeness requirements, compute means, compute percentiles, compute DV’s , and generate raw and summary db’s for complete sites only (‘coarse comp final.sas’).

Construction of PM_{10} databases

For SP analyses, two PM_{10} databases were created and used. Both versions relied on daily summary AQS extractions; SAS code (‘airs_dailysum_pm10dv.sas’) was used for the extraction. The AQS daily summaries table encompasses 24-hour filter measurements and hourly data aggregated to a 24-hour basis (as noted above in the $PM_{10-2.5}$ database discussion). Of the latter type, only the valid ones (those with DAILY_CRITERIA_IND=’Y’, signifying 18+ hourly observations per day) were used. The difference between the two created PM_{10} databases was in the treatment of event-flagged data. One version (in, EDT_EDT_ID) was only used to calculate the county NAAQS exceedance information (counts/population/percentage); see Analysis 10 below.

Note that this second database corresponds exactly to what was used to generate the official 2001-2003 PM_{10} Design Values.

Boxplot Figures

Many of the generated 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.
- If shown, the distribution maximum and minimum are shown as asterisks.

Analysis 1 – Summaries and boxplots of $PM_{2.5}$ and $PM_{10\text{-}2.5}$ annual mean and 98th percentile DV's, by region

Goals:

? To characterize the typical average concentration levels of PM_{10} and $PM_{2.5}$ for different U.S. regions.

? To make comparisons of the size cuts compare.

Outputs:

- o Summary statistics were generated by region. See tables in Output A.1a.
- o Boxplots were generated of the distribution of site-level annual means and 98th percentile by region. See Output A.1b.

Methods:

- The SAS procedure UNIVARIATE was used to generate the summary statistics.
- SAS code ('inputbox mean 98p.sas' and 'boxplot pmf pmc.sas') was used generate the boxplots.

Analysis 2 – Maps of $PM_{2.5}$, $PM_{10\text{-}2.5}$, and PM_{10} county maximum annual mean and 98th percentile DV's, by region

Goals:

? To identify specific geographic areas with high and low annual mean and 98th percentile concentration levels.

Outputs:

- o $PM_{2.5}$ maps are shown in Output A.2a.
- o $PM_{10\text{-}2.5}$ maps are shown in Output A.2b.
- o PM_{10} maps are shown in Output A.2c.

Methods:

- Each county (with a complete site) was assigned the value of the site with the highest stated statistic (annual mean or 98th percentile DV).
- SAS code, 'map4shade.sas', was used to generate the $PM_{2.5}$ and $PM_{10\text{-}2.5}$ maps.

- SAS code, ‘bwfammap.sas’ and ‘bwcntymap2.sas’, was used to generate the PM_{10} maps.

Analysis 3 - Event-flagged data, $PM_{2.5}$ and $PM_{10-2.5}$

Goals:

- To identify the types of events that are flagged in AQS.
- To determine if there are significant amounts of event-flagged PM data.
- To determine if ‘high’ sites flag more data than ‘low’ sites.
- To see if events impact DV’s.
- $PM_{2.5}$: To ascertain whether any DV’s changed from ‘violating the standard’ to ‘meeting the standard’ after removing their event-flagged data
- To see if the impacts are different for ‘high’ versus ‘low’ sites
- To determine whether data distributions are similar for sites that flag data compared to those that do not flag data.
- $PM_{2.5}$: To evaluate the specific impact of episodic events on NAAQS-type metrics (case studies).

Outputs:

- Various tables, plots, and related discussion; see Output 3a for $PM_{2.5}$ and Output 3b for $PM_{10-2.5}$.

Methods:

- Subsequent to $PM_{2.5}$ and $PM_{10-2.5}$ ‘regular’ (all data included) mean and percentile processing (as noted under general/background), SAS code was used to recalculate the indicator metrics excluding event-flagged data.
- Only exceptional- and natural-event flagged data were ignored. This includes the following flag situations: ‘A’ (high winds); ‘C’ (volcanic eruptions); ‘E’ (forest fire); ‘F’ (structural fire); ‘J’ (construction/demolition); ‘K’ (agricultural tilling); ‘L’ (highway construction); ‘M’ (rerouting of traffic); ‘O’ (infrequent large gatherings); ‘P’ (roofing operations); ‘Q’ (prescribed burning); ‘R’ (clean up after a major disaster) ‘S’ (seismic activity); and ‘U’ (Sahara dust).
- Unlike in production design value (PDV) processing for $PM_{2.5}$ and PM_{10}, the AQS regional concurrence indicator was not evaluated. Thus, the concurrence being set to ‘yes’ was <u>not</u> a requisite for flagged data to be excluded.
- For $PM_{10-2.5}$, a site-day estimate was assumed to be event-flagged if either the PM_{10} or the $PM_{2.5}$ data were event- flagged
- SAS code (‘ex events fine.sas’ and ‘quebec.sas’) was used to evaluate the $PM_{2.5}$ events.
- SAS code (‘ex events coarse.sas) was used to evaluate the $PM_{10-2.5}$ events.

Analysis 4 - Comparisons of site-level annual means to 98th percentiles, $PM_{2.5}$ and $PM_{10-2.5}$

Goals:

- To evaluate the relationship between site-level annual means and site-level 98th percentiles.

Outputs:

- See Output A.4.

Methods:

- The distributions of site-level 98th percentiles were plotted by intervals of site-level mean levels.
- SAS code was used to generate the $PM_{2.5}$ and $PM_{10-2.5}$ plots (‘pmf boxplot p98 intmean.sas’ and ‘pmc boxplot p98 intmean.sas’).

Analysis 5 – Regional correlations of $PM_{2.5}$, $PM_{10-2.5}$, and PM10

Goals:

? To evaluate the correlation among the three size cuts.

Outputs:

- See Output A.5.

Methods:

- Because the represented periods are different for $PM_{2.5}$, $PM_{10-2.5}$, and PM_{10} (e.g., For $PM_{10-2.5}$, the most recent 12, 8, or 4 quarters were utilized; for PM_{10} and $PM_{2.5}$, all 12 quarters were needed) and also because completeness was applied independently, the selected time periods did not necessarily match. If the common time periods of constituent raw data (for the sites that met the parameter selection criteria) were used for this analysis, some sites common to multiple 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. 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).
- 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 then averaged by region.
- SAS code was used to calculate the correlations and produce the plot (‘procorr.sas’).

Analysis 6 – Distribution of ratios of 24-hour average $PM_{2.5}$ to PM10, by region

Goals:

? To identify typical site average 24-hour ratios of $PM_{2.5}$ to PM10, by region.

Outputs:

- See Output A.6.

Methods:

- The ratio of $PM_{2.5}$ to PM_{10} was first calculated for each site-day. The site-day ratios of $PM_{2.5}$ to PM_{10} were then averaged by site and the distribution of the site ratios plotted by region.
- SAS code (‘ratio of pmf to pmt.sas’) was used for the analysis.

Analysis 7 – $PM_{10-2.5}$ equivalence to PM_{10} NAAQS (daily standard)

Goals:

? To evaluate how equivalent (well correlated) various $PM_{10-2.5}$ design value type metrics are to the existing PM_{10} 24-hour standard?
? To see if the relationships are different for 'high' PM_{10} levels?
? To ascertain if the relationships vary across regions?
? To estimate the levels for various $PM_{10-2.5}$ design value type metrics that would correspond to the 150 Ug/m3 level for the current PM_{10} 24-hour standard.

Outputs:

- See Output A.7.

Methods:

- For the same reason noted in the first bullet for 'Analysis 5', the $PM_{10-2.5}$ database was utilized for this evaluation. Hence, comparisons of PM_{10} and $PM_{10-2.5}$ metrics represented the same timeframe (and same days).
- The recalculated $PM_{10-2.5}$ percentile metrics (excluding event flagged data, per Analysis 3) were used in this evaluation.
- For perspective, the actual, official PM_{10} DV's (i.e., emanating from the db excluding concurred event data) were also merged with data and evaluated against the $PM_{10-2.5}$ metrics.
- SAS code ('pmc equivalence to pmt.sas') was used for the analysis.

Analysis 8 – Evaluation of spatial averaging (SA), $PM_{2.5}$

Goals:

? To determine if spatial averaging allows large populations to go unprotected (i.e., to estimate the population in areas that could possibly use SA (utilizing current criteria)).
? To determine if there are large differences between 'regular' DV's (based on highest site in area) and DV's calculated with SA.
? To evaluate potential concerns with spatial averaging. (for example, are the would-be violating sites that could utilize SA located in lower-income, high percentage-minority, and/or lower education area (based on Census tract information) than the overall area?)
? To evaluate the current criteria for using SA. To opine the appropriateness of revising the criteria.

Outputs:

- See Output A.8.

Methods:

- Initially started with the default SP $PM_{2.5}$ database (all sites with 11+ samples in each of the 12 quarters 2001-2003). Eliminated sites that are not (officially) compared to annual standard. (AQS Site ID's: '180890022', '180970066', '180970043', '170311016', '171190023', '170990007', '440070020', '481410053', and '291250001')
- Initially enforced the CFR spatial criteria of: 1) 0.6 overall correlation between sites, and 2) no more than 20% difference in site annual mean and spatial annual mean. The criterion that all SA sites should be impacted by similar emissions was not evaluated.

- Enforced CFR data handling requirement that if SA annual mean is less than or equal to the annual standard, then only SA sites with 75%+ capture each of the 4 Q's would have their annual mean included in the spatial annual average. (Only 11+ samples required in each of the 4 Q's if the spatial annual mean was greater than the evaluated annual standard.) Changed level of standard (and completeness check) from 15 to 14 for accurate evaluation of SA effect on those standard levels.
- For 'area' definitions, utilized OMB definitions for Core-Based Statistical Areas (CSA's) and Combined Statistical Areas (CSA's). If multiple sites were not located in a defined area, then area was assumed to be the county.
- Constructed SA set of sites by initially considering all sites in the area. If a site-pair correlation was less than cutoff, the lower DV site was eliminated. If a remaining set did not meet annual mean difference criterion, then the lowest DV site was omitted from the set and the revised set tried. This continued until the reduced set of sites met criteria, or until less than 2 sites were left. Note: Undoubtedly, different combinations of sites (selected with some rationale and/or at random) could/would meet criteria and yield different results.
- Only considered (for SA) areas with 1) a regular DV greater than the evaluated annual standard level, and 2) a spatial DV greater than any (valid) non-SA site DV in the area.
- Evaluated appropriateness of 0.6 (correlation) and 20% (max difference in annual means) levels by comparing to typical universe values.
- Tightened the correlation criterion to 0.9 and the annual mean difference criterion to 10% to evaluate changes in results. SAS code ('spatial avg.sas') was used for the analysis.
- SAS code ('spatial avg.sas' and 'simple spatial.sas') was used to conduct the analyses.

Analysis 9 – Evaluation of 'high' $PM_{2.5}$ values

Goals:

? To identify the minimum number of days per year a site is permitted to exceed the annual 98th, 99th, and other percentiles.

? To evaluate the (entire) daily distributions of data plotted by 98th (and 99th) percentile-level intervals.

? To evaluate the daily distributions of data exceeding site-level 98th (and 99th) DV's plotted by 98th (and 99th) percentile intervals.

? To ascertain the actual number and percentage of days (site average, minimum, & maximum), for the 3-year period 2001-2003, where the concentration was significantly above the site 98th or 99th percentiles. [Significant defined as 5+ $\mu g/m^3$.]

Outputs:

o See Output A.9.

Methods:

- SAS code ('dist above p98.sas') was used for the analysis.

Analysis 10 – Estimated percentage of counties not likely to meet alternative $PM_{2.5}$ and $PM_{10-2.5}$ standards and existing PM_{10} NAAQS

Goals:
- ? To estimate the number, percentage and population of counties on a National level not likely to meet alternative PM standards.
- ? To estimate the percentage of counties on a regional basis not likely to meet alternative PM standards.

Outputs:
- o See Output A.10.

Methods:
- Recalculated $PM_{2.5}$ annual and percentile metrics *excluding event-flagged data* and were used in this evaluation. However, unlike Analysis 3, which excluded *all* event-flagged data, this particular analysis only excluded *regionally-concurred* (per AQS) event-flagged data (just like in official DV's).
- Additionally, for the annual $PM_{2.5}$ standard level evaluation (by itself, and in tandem with a daily standard), the sites officially exempted from the annual standard (AQS Site ID's): '180890022', '180970066', '180970043', '170311016', '171190023', '170990007', '440070020', '481410053', and '291250001') were not considered to be in violation of the annual standard no matter the level. Essentially, the annual mean DV was set to zero for these sites. These sites *were* compared to the alternative daily standard levels.
- As noted in the background $PM_{10-2.5}$ database construction section, the 'extended' database was used for this analysis instead of the 'regular' database. The extended db includes the 'regular' db plus data pairs from non-collocated (but nearby - up to 5 miles away) FRM/FEM sites. The $PM_{10-2.5}$ estimate was anchored at the PM_{10} site. The assumption is that $PM_{2.5}$ is fairly homogenous, but PM_{10} is not. [The rationale for expanding the $PM_{10-2.5}$ db to included non-collocated pairs of data is as follows: Many 'high' PM_{10} sites do not have collocated $PM_{2.5}$ because of disparate monitoring objectives. For PM_{10} the central objective is 'highest concentration'; for $PM_{2.5}$ the main NAAQS objective is 'population exposure'. Hence, by not doing including these non-collocated pair, we would be ignoring many potentially high $PM_{10-2.5}$ locations.] Several PM_{10} sites identified as source-oriented and not also population exposure were omitted from the extended database because it was felt that they were not likely candidates for a future $PM_{10-2.5}$ network. These sites, identified by EPA regional staff, are (AQS Site ID's): '090090018', '290970003', '295100092', '401010167', '440070020', '450430006', '450630009', '560050874', '560050885', '560050891', '560050894', and '560050907'. 712 sites (located in 382 counties) are in the $PM_{10-2.5}$ 'extended' database.
- Recalculated $PM_{10-2.5}$ percentile metrics *excluding event-flagged data* were used in this evaluation. For $PM_{10-2.5}$, AQS regional concurrence was not evaluated (i.e., same process as used in Analysis 3, albeit to the bigger db).
- SAS code ('whatif county counts pmf.sas', 'whatif county counts pmc.sas', and 'counties_not_meeting_naaqs.sas') was used for the analyses.

Analysis 11 – Monthly patterns of urban $PM_{2.5}$ and $PM_{10-2.5}$

Goals:

? To identify monthly patterns, by region, in concentrations of $PM_{2.5}$ and $PM_{10-2.5}$

Outputs:

- $PM_{2.5}$ boxplots are plots are shown in Output A.11a. $PM_{10-2.5}$ boxplots are shown in Output A.11b.

Methods:

- Only data from monitors with AQS 'location setting' of 'URBAN AND CENTER CITY' or 'SUBURBAN' were used. Hence, the term 'urban' actually encompasses 'suburban' sites as well.
- All 24-hour average values (for complete 'urban' sites from the 'regular' $PM_{2.5}$ and $PM_{10-2.5}$ db's) were averaged together by region-month.
- In these boxplots, 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. Whiskers (95^{th} and 5^{th} percentiles) were not plotted.

PM2.5: Summary Statistics for Site-Level Annual Mean DV

PMREG	PMREGDE	n	mean	max	p95	p75	median	p25	p05	min
0	Not in PMF	17	6.629412	11.9	11.9	7.4	6.4	5.1	3.9	3.9
1	Northeast	121	13.20248	17.3	16.4	14.6	13.3	12	9.6	6.5
2	Southeast	216	12.52407	18	15.7	13.9	12.55	11.2	9.1	7.4
3	Industrial M	217	14.60461	21.2	17.4	15.7	14.7	13.5	11.4	6.6
4	Upper Midv	71	9.988732	13.9	12.6	11.3	10.5	9	6	5.5
5	Southwest	33	8.515152	16.9	14.4	10.7	7.8	6.6	4.4	4
6	Northwest	110	9.37	17	13.4	10.8	9.1	7.8	5.6	4.5
7	Southern C	42	16.63333	27.8	25.2	21.3	16.9	12	6.9	6.2
	U.S.	827	12.45961	27.8	17.2	14.6	12.6	10.4	6.6	3.9

PM2.5: Summary Statistics for Site-Level 98th Percentile

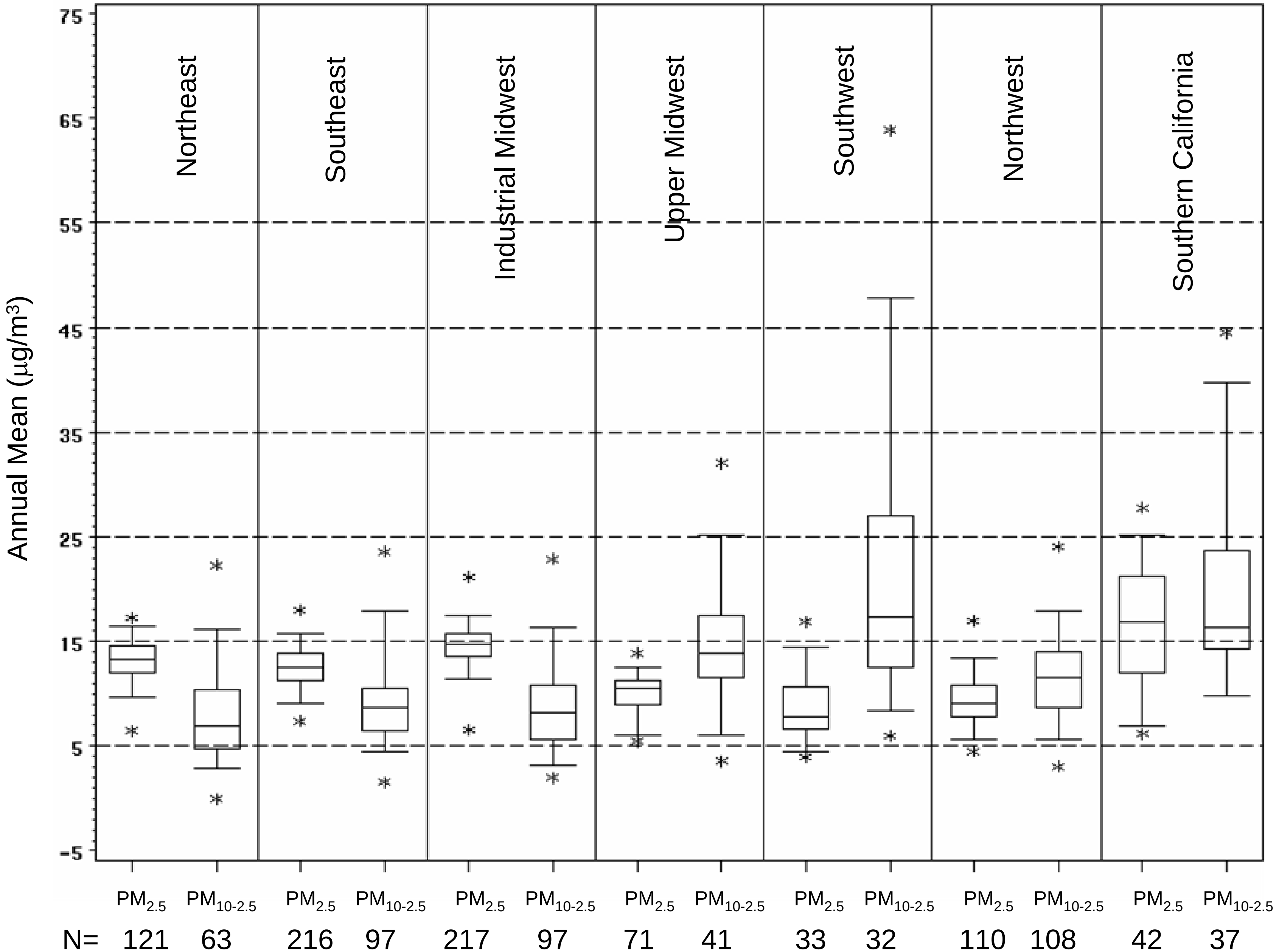
PMREG	PMREGDE	n	mean	max	p95	p75	median	p25	p05	min
0	Not in PMF	17	16.47059	40	40	17	15	13	9	9
1	Northeast	121	36.3719	48	43	40	37	33	29	25
2	Southeast	216	28.5787	40	36	31	29	26	20	17
3	Industrial M	217	36.29954	63	43	39	36	34	29	18
4	Upper Midv	71	25.08451	33	30	28	26	23	16	14
5	Southwest	33	22.09091	49	46	28	20	16	10	10
6	Northwest	110	31.87273	62	54	40	30.5	23	15	11
7	Southern C	42	45.38095	76	72	62	45.5	29	20	18
	U.S.	827	32.22854	76	46	37	32	27	17	9

PM10-2.5: Summary Statistics for Site-Level Annual Mean DV

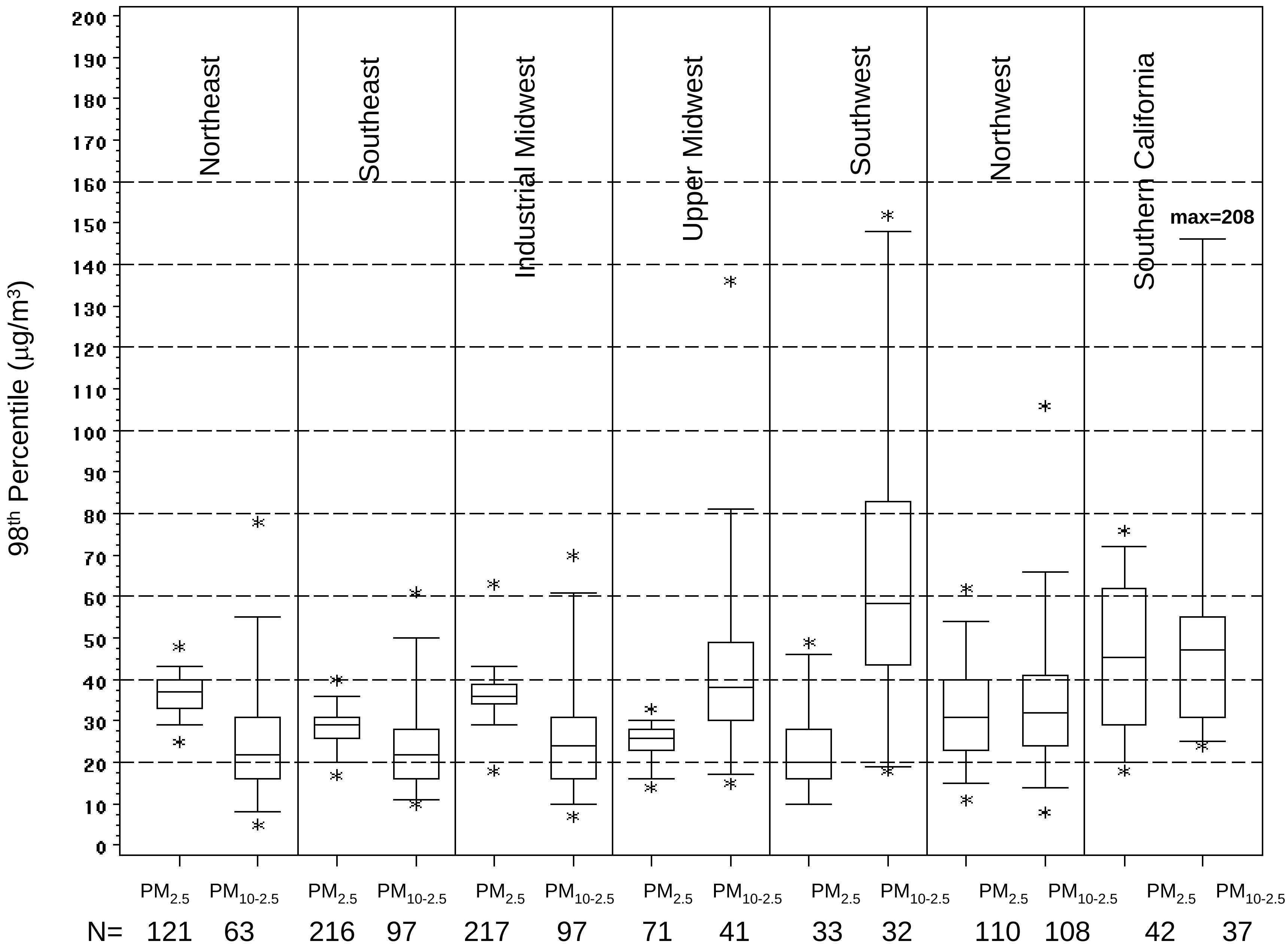
PMREG	PMREGDE	n	mean	max	p95	p75	median	p25	p05	min
0	Not in PMR	14	16.75259	30.2	30.2	24.6	15.0	11.0	1.8	1.8
1	Northeast	63	7.877622	22.3	16.2	10.4	6.9	4.8	2.8	0.0
2	Southeast	97	9.311192	23.6	17.9	10.6	8.7	6.5	4.5	1.6
3	Industrial M	97	8.842588	22.9	16.3	10.8	8.2	5.6	3.1	2.0
4	Upper Midw	41	14.37395	32.1	25.2	17.5	13.8	11.6	6.1	3.6
5	Southwest	32	21.1939	63.9	47.8	26.9	17.3	12.6	8.3	6.0
6	Northwest	108	11.58091	24.1	17.9	14.0	11.6	8.6	5.6	3.1
7	Southern C	37	19.80212	44.5	39.8	23.7	16.3	14.3	9.8	9.8
	U.S.	489	11.74376	63.9	24.9	14.7	10.5	7.0	4.1	0.0

PM10-2.5: Summary Statistics for Site-Level 98th Percentile

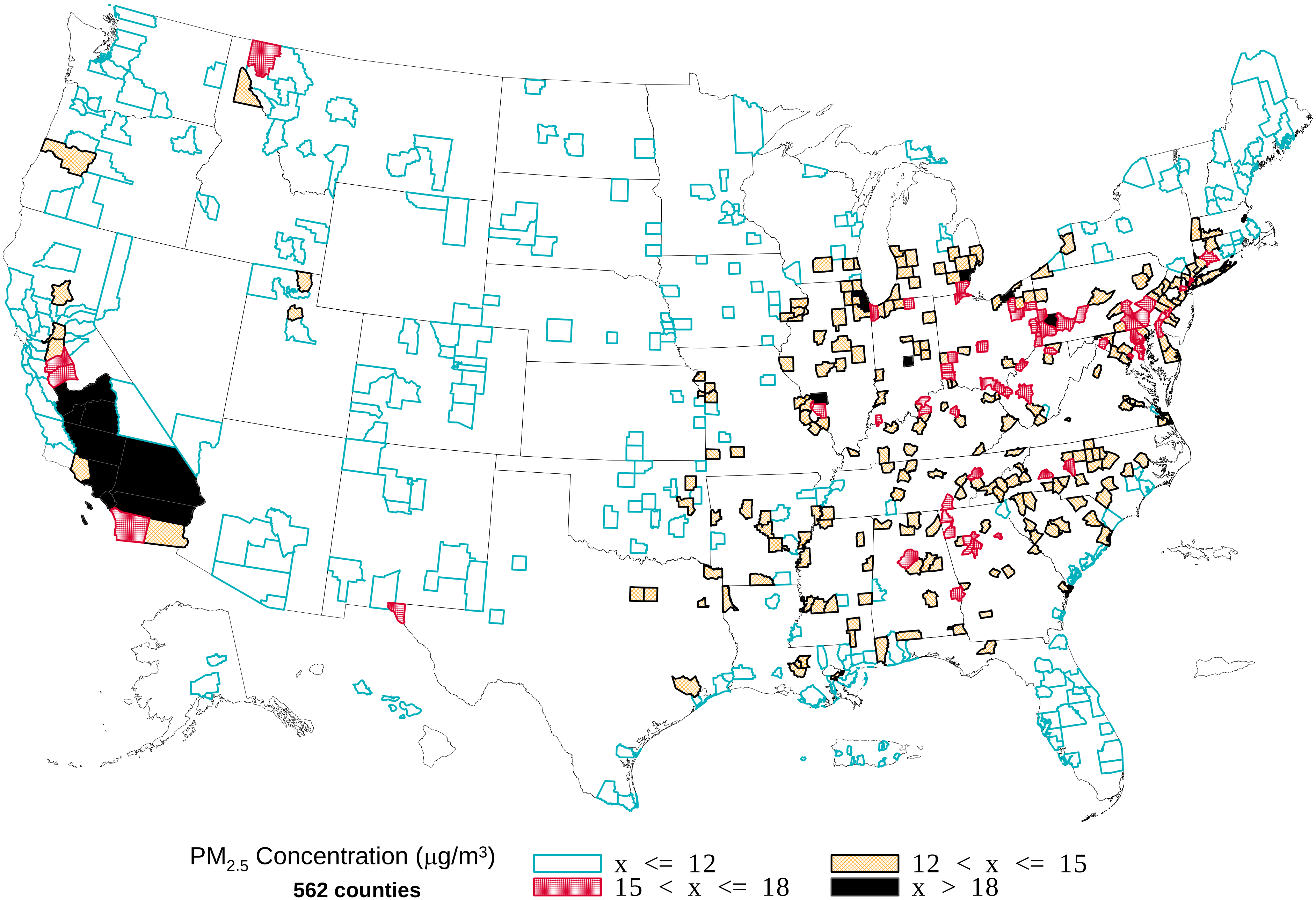
PMREG	PMREGDE	n	mean	max	p95	p75	median	p25	p05	min
0	Not in PMF	14	48	89	89	67	50.5	22	10	10
1	Northeast	63	25.39683	78	55	31	22	16	8	5
2	Southeast	97	24.5567	61	50	28	22	16	11	10
3	Industrial M	97	25.47423	70	61	31	24	16	10	7
4	Upper Midw	41	42.4878	136	81	49	38	30	17	15
5	Southwest	32	64.75	152	148	83	58.5	43.5	19	18
6	Northwest	108	33.88889	106	66	41	32	24	14	8
7	Southern C	37	52.97297	208	146	55	47	31	25	24
	U.S.	489	33.86299	208	74	41	28	20	11	5



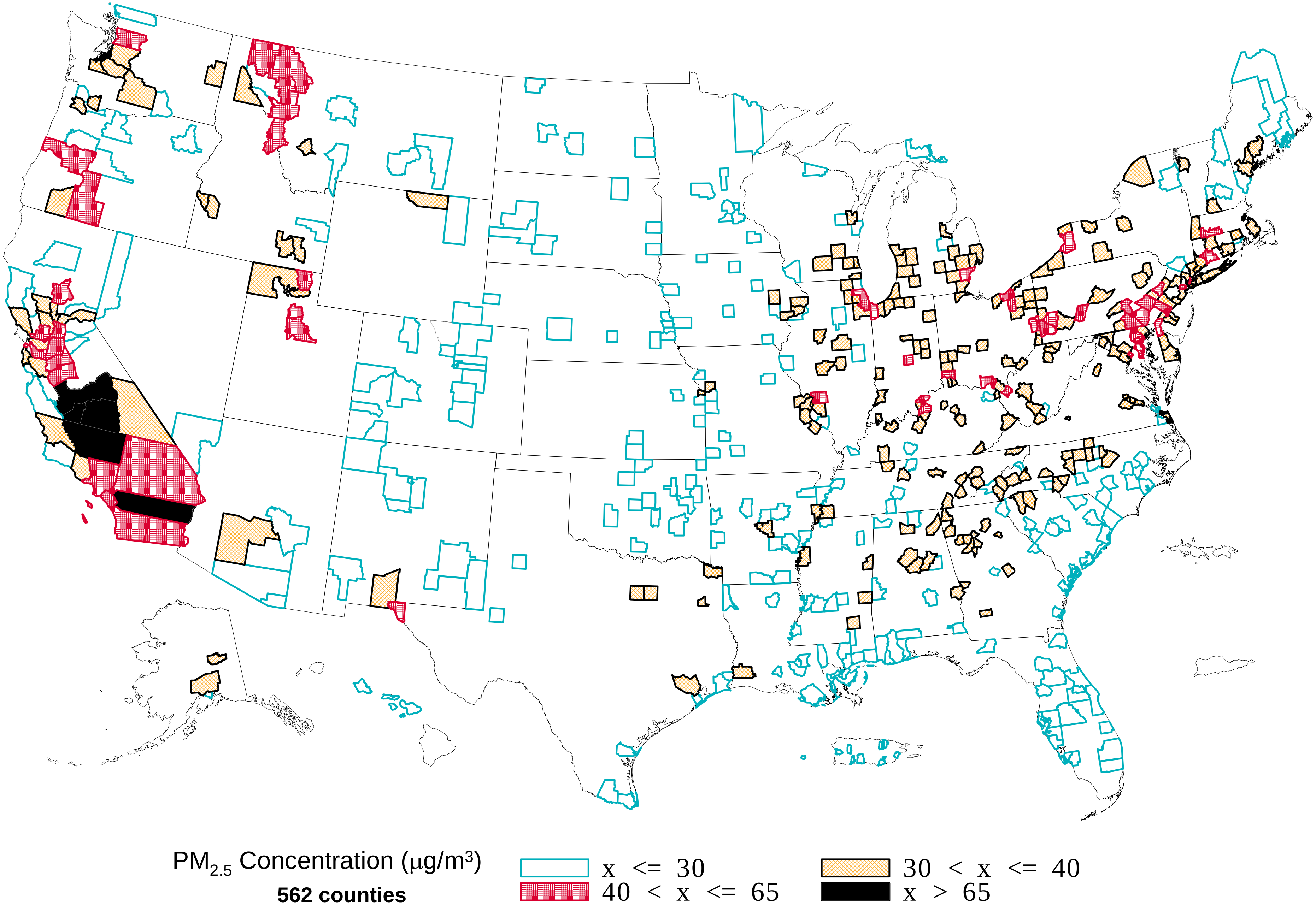
Distribution of annual mean $PM_{2.5}$ and estimated annual mean $PM_{10\text{-}2.5}$ concentrations by region, 2001-2003.
N = number of sites.



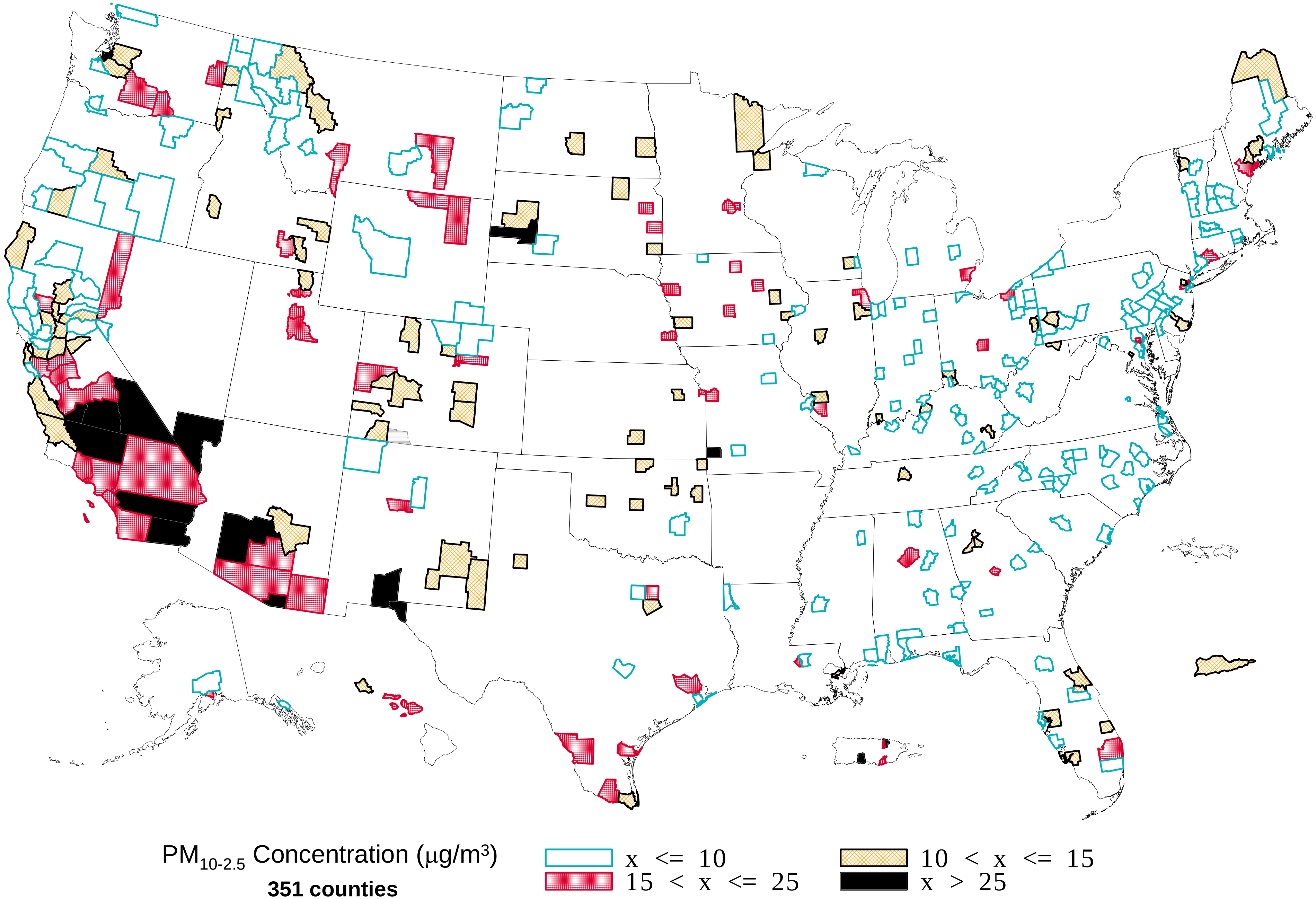
Distribution of 98th percentile 24-hour average $PM_{2.5}$ and estimated $PM_{10-2.5}$ concentrations by region, 2001-2003.
N = number of sites.



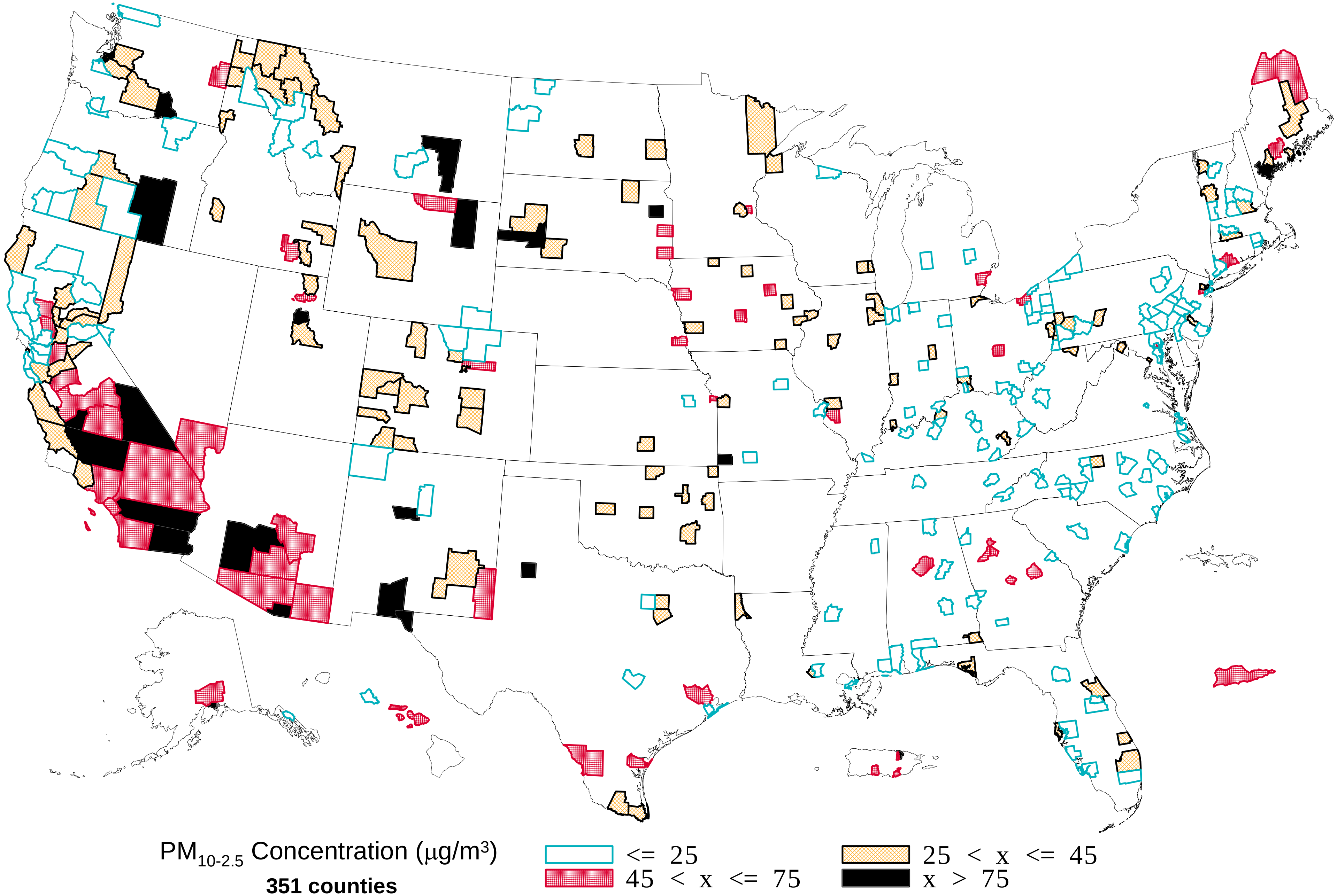
County-level maximum annual mean $PM_{2.5}$ concentrations, 2001-2003.



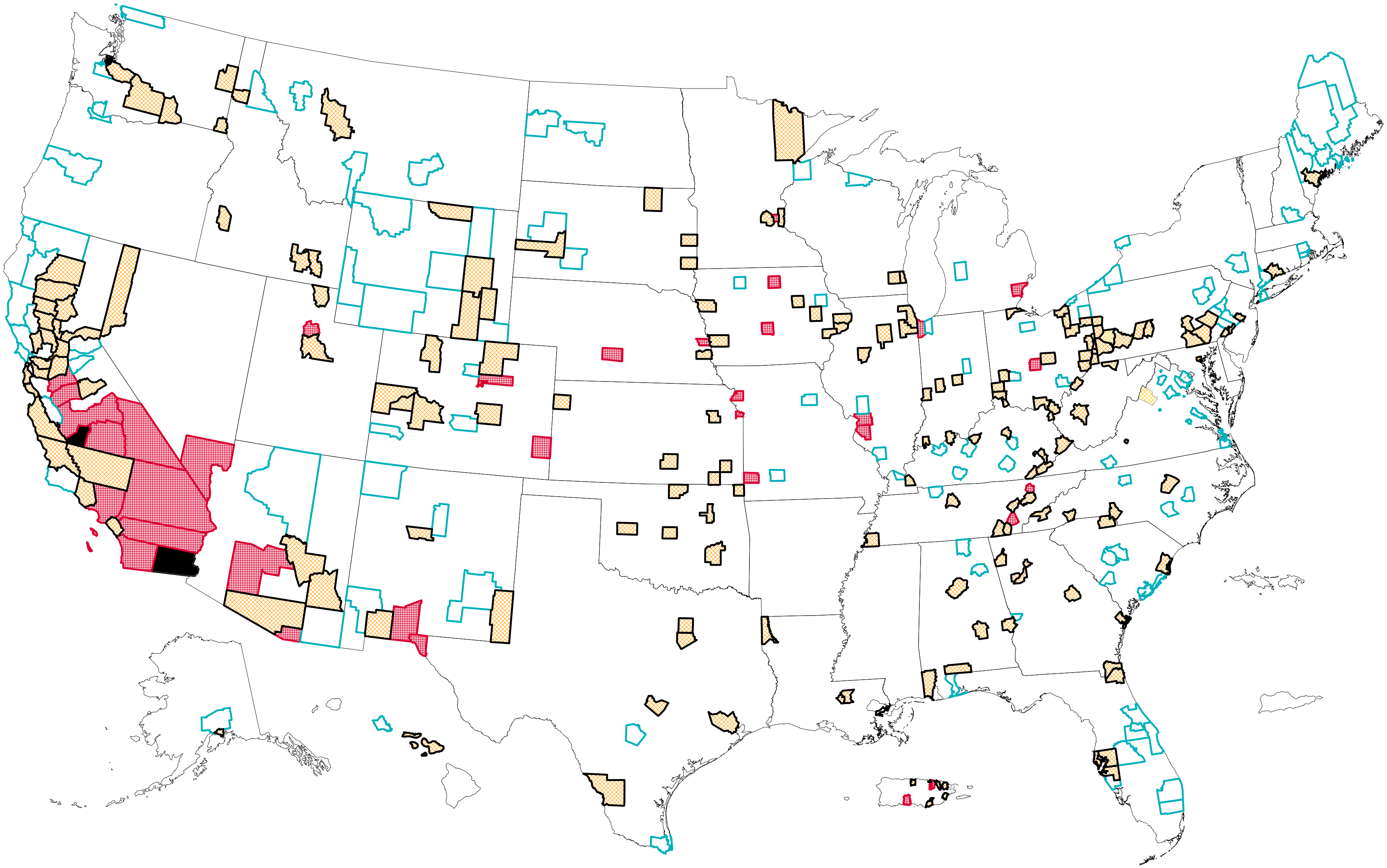
County-level maximum 98th percentile 24-hour average $PM_{2.5}$ concentrations, 2001-2003.



Estimated county-level maximum annual mean $PM_{10\text{-}2.5}$ concentrations, 2001-2003.



Estimated county-level maximum 98th percentile 24-hour average $PM_{10\text{-}2.5}$ concentrations, 2001-2003.

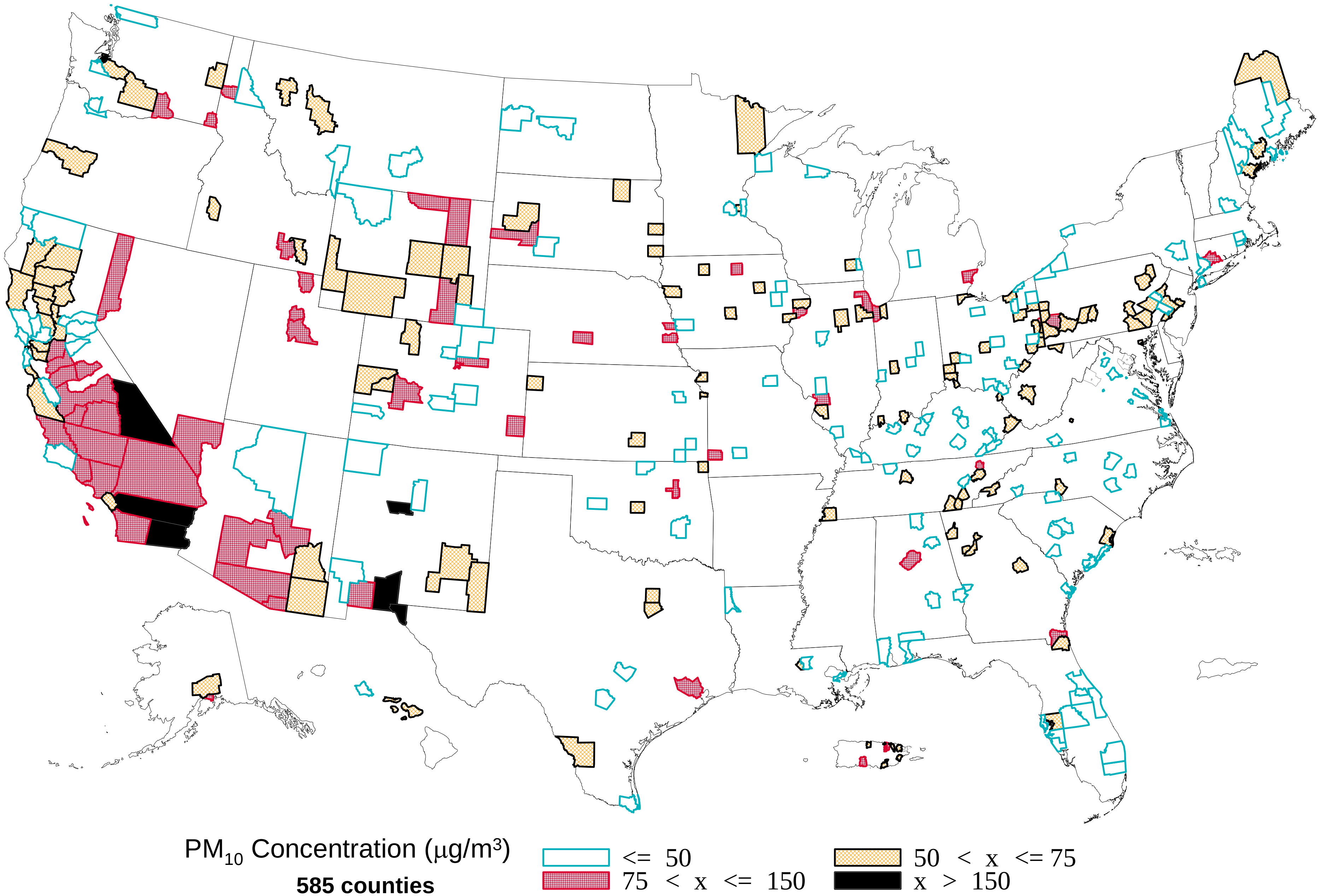


PM$_{10}$ Concentration (µg/m^3)

585 counties

<= 20 | 20 < x <= 30 | 30 < x <= 50 | x > 50

County-level maximum annual mean PM$_{10}$ concentrations, 2001-2003.



County-level maximum 98th percentile 24-hour average PM_{10} concentrations, 2001-2003.

Episodic Events – PM2.5

- Questions:
  1. What types of events are flagged in AQS?
  2. Are there a significant amount of event-flagged data in AQS?
  3. Do ‘high’ sites (> 15 annual DV) flag more data than ‘low’ sites (≤ 15)?
  4. How do events impact DV’s?
  5. Did any DV’s go from ‘violates the std’ to ‘meets the std’
  6. Are the impacts different for ‘high’ vs ‘low’ sites?
  7. Are data distributions similar for sites that flag data vs. sites that don’t flag data?
  8. What is specific impact in select areas (case studies)

1. What types of events are flagged in AQS?

PM2.5 flag counts, 2001-2003 - All data (meets or not meets completeness)

Flag Description	Event Class	Flag Count	Percent of Event Flags	Number of Sites Reporting
Forest Fire	natural	661	39.9%	273
Sahara Dust	natural	306	18.5%	12
Construction/Demolition	except.	253	15.3%	13
Highway Construction	except.	133	8.0%	4
Volcanic Eruptions	natural	99	6.0%	10
Roofing Operations	except.	51	3.1%	5
Structural Fire	except.	29	1.8%	4
Clean Up After A Major Disaster	except.	29	1.8%	24
Rerouting Of Traffic	except.	27	1.6%	1
High Winds	natural	25	1.5%	14
Infrequent Large Gatherings	except.	17	1.0%	6
Prescribed Burning	except.	17	1.0%	10
Agricultural Tilling	except.	8	0.5%	4
Seismic Activity	natural	1	0.1%	1
Total		1,656		339

PM2.5 Flag Counts, 2001-2003 - Data for sites that meet completeness

Flag Description	Event Class	Flag Count	Percent of Event Flags	Number of Sites Reporting
Forest Fire	natural	490	34.6%	197
Sahara Dust	natural	296	20.9%	10
Construction/Demolition	except.	239	16.9%	10
Highway Construction	except.	133	9.4%	4
Volcanic Eruptions	natural	99	7.0%	10
Roofing Operations	except.	51	3.6%	5
Rerouting Of Traffic	except.	27	1.9%	1
Clean Up After A Major Disaster	except.	26	1.8%	21
High Winds	natural	18	1.3%	10
Infrequent Large Gatherings	except.	14	1.0%	4
Structural Fire	except.	9	0.6%	2
Prescribed Burning	except.	8	0.6%	5
Agricultural Tilling	except.	6	0.4%	2
Seismic Activity	natural	1	0.1%	1
Total		1,417		249

- Most flagged data relate to natural events (~64%)
- Forest fires is the most common event flagged (looking by flag or by site)
- 30% of all complete sites (249 / 827) flagged at least one concentration
- For data from all sites (complete or not): 6 of the 10 highest PM2.5 values were flagged; 41 of the top 100; and 194 of the top 1000
- For data from complete sites: 7 of the 10 highest PM2.5 values were flagged; 39 of the top 100; and 148 of the top 1000

***Complete sites defined as those with 12 quarters of 11+ samples.**

2. Are there a significant amount of event-flagged data in AQS?

Percentage of event-flagged data - for complete sites (827)

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	0.3%	2.6%	3.2%	3.8%	4.8%	7.2%

All complete sites…. But some sites may not flag. (Hence, perhaps biased low). Reference as 'complete sites'

Percentage of event-flagged data - at sites with at least 1 flagged point [249 sites]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	1.1%	8.8%	10.5%	12.7%	15.8%	23.8%

Complete sites with at least one event flag 2001-2003. But some other sites may have flagged if event occurred. (Hence, perhaps biased high). Reference as 'flag sites'

Percentage of event-flagged data at complete sites where RO has at least 1 flagged datapoint (not necessarily at all sites) [423 sites]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	0.7%	5.2%	6.2%	7.5%	9.3%	14.0%

Complete sites where the reporting organization (RO) has flagged at least one event 2001-2003, though not necessarily at all sites. Reference as 'RO flaggers'. [The RO knows how to flaf.]

- Event-flagged data only account for .3% to 1.1% of all reported observations
- However, they account for considerably higher percentage of high values (i.e., values $\geq$ 95th, 96th, 97th, 98th, 99th percentile) ~ 5 %– 14%

3. Do ‘high’ sites (> 15 annual DV) flag more data than ‘low’ sites (< 15)?

Percentage of event-flagged data at complete sites where RO has at least 1 flagged datapoint (not necessarily at all sites) [423 sites]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	0.7%	5.2%	6.2%	7.5%	9.3%	14.0%

Same as previous page (bottom). Break out by high / low.

Percentage of event-flagged data - RO flaggers, sites > 15.0 [58 sites]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	1.0%	4.3%	5.2%	6.8%	8.3%	12.3%

High sites

Percentage of event-flagged data - RO flaggers, sites ≤ 15.0 [365 sites]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	0.6%	5.3%	6.4%	7.6%	9.5%	14.3%

Low sites

- Not much difference

4. How do events impact DV's?

Reductions (µg/m3) in annual and 24-hour design values as a result of exempting event-flagged data - complete sites [827 sites]

	Reduction (ug/m3) in Stated Metric					
Site change	Annual DV	95th Percentile	96th Percentile	97th Percentile	98th Percentile	99th Percentile
Maximum	1.5	9	9	15	18	21
95th percentile	0.2	1	1	1	2	4
75th Percentile	0	0	0	0	0	0
Average	0.04	0.19	0.22	0.28	0.39	0.64
Median	0	0	0	0	0	0
25th percentile	0	0	0	0	0	0
5th Percentile	0	0	0	0	0	0
Minimum	-1.8	0	0	0	0	0

Reductions (µg/m3) in annual and 24-hour design values as a result of exempting event-flagged data - flag sites [249 sites]

	Reduction (ug/m3) in Stated Metric					
Site change	Annual DV	95th Percentile	96th Percentile	97th Percentile	98th Percentile	99th Percentile
Maximum	1.5	9	9	15	18	21
95th percentile	0.4	2	3	3	4	8
75th Percentile	0.2	1	1	1	2	3
Average	0.15	0.62	0.74	0.94	1.30	2.12
Median	0.1	0	0	0	1	1
25th percentile	0.1	0	0	0	0	0
5th Percentile	0	0	0	0	0	0
Minimum	-1.8	0	0	0	0	0

Reductions (µg/m3) in annual and 24-hour design values as a result of exempting event-flagged data - RO flaggers [423 sites]

	Reduction (ug/m3) in Stated Metric					
Site change	Annual DV	95th Percentile	96th Percentile	97th Percentile	98th Percentile	99th Percentile
Maximum	1.5	9	9	15	18	21
95th percentile	0.3	2	2	3	3	6
75th Percentile	0.1	0	1	1	1	1
Average	0.09	0.37	0.43	0.56	0.77	1.25
Median	0	0	0	0	0	0
25th percentile	0	0	0	0	0	0
5th Percentile	0	0	0	0	0	0
Minimum	-1.8	0	0	0	0	0

- The last table (RO flaggers) probably represents the best guess at national average effect.
- On average, removing flagged data would reduce annual DV's by about .1 ug/m3, 98^{th} percentiles by about .8 ug/m3, and 99^{th} percentiles by 1.3 ug/m3
- 25%+ sites would have 1 ug/m3 lower percentiles (96^{th}-99^{th}) if flagged data were omitted

5. Did any DV's go from 'violates the std' to 'meets the std'

- Three complete sites that violate the annual std of 15.0 would meet the std if event-flagged data were excluded
- However, in all 3 situations there exists additional sites that violate the std with or without event data (with much higher DV's)

site	DV all data	DV minus flags	state_name	county_name	csa_name
100031012	15.2	15.0	Delaware	New Castle	Philadelphia-Camden-Vineland, PA-NJ-DE-MD
245100007	15.1	15.0	Maryland	Baltimore (City)	Washington-Baltimore-Northern Virginia, DC-MD-VA-WV
470654002	15.2	15.0	Tennessee	Hamilton	Chattanooga-Cleveland-Athens, TN-GA

- No sites violate the daily std of 65 but would meet it if flagged data were excluded.

6. Are the impacts different for 'high' vs 'low' sites?

Reductions (µg/m3) in annual and 24-hour design values as a result of exempting event-flagged data - RO flaggers, sites > 15.0 [58 sites]

	Reduction (ug/m3) in Stated Metric					
Site change	Annual DV	95th Percentile	96th Percentile	97th Percentile	98th Percentile	99th Percentile
Maximum	0.3	3	3	3	4	7
95th percentile						
75th percentile						
Average	0.05	0.47	0.45	0.47	0.83	1.26
Median	0.1	0	0	0	0	0
25th percentile						
5th Percentile						
Minimum	-1.8	0	0	0	0	0

Reductions (µg/m3) in annual and 24-hour design values as a result of exempting event-flagged data - RO flaggers, sites < 15.0 [365 sites]

	Reduction (ug/m3) in Stated Metric					
Site change	Annual DV	95th Percentile	96th Percentile	97th Percentile	98th Percentile	99th Percentile
Maximum	1.5	9	9	15	18	21
95th percentile						
75th percentile						
Average	0.09	0.35	0.43	0.57	0.76	1.25
Median	0	0	0	0	0	0
25th percentile						
5th Percentile						
Minimum	0	0	0	0	0	0

- Not much difference in effect on percentiles DV's
- High sites have a about double the reduction in annual DV's… but still small effect on average (less than .1 ug/m3)
- Some sites have considerable effects

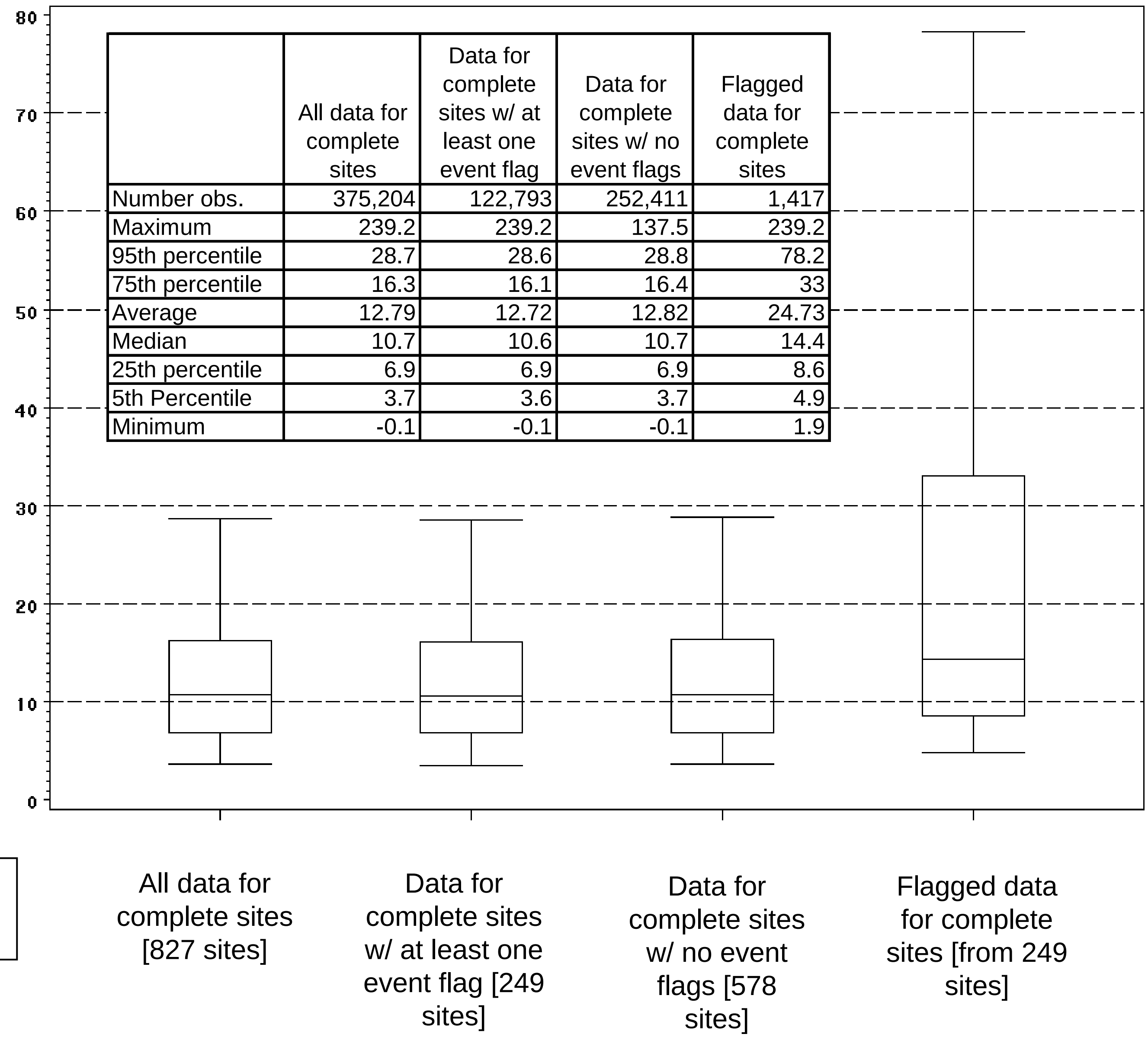
7. Are data distributions similar for sites that flag data vs. sites that don't flag data?

- See next 2 slides
- 2nd slide more accurate comparison
- Not much difference in distributions for flag sites vs. no flag sites.
- But, there are obvious differences in data distributions of all data vs. flagged data.
  - Flagged data generally higher, average concentrations 12.5 - 12.8 for all data (at comp sites) vs. 24.8 for flagged data

7. Are data distributions similar for sites that flag data vs. sites that don't flag data?

Distribution of PM2.5 concentrations: All data at complete sites, data for complete sites w/ event flags, data for complete sites w/out flags, flagged data from complete sites

	All data for complete sites	Data for complete sites w/ at least one event flag	Data for complete sites w/ no event flags	Flagged data for complete sites
Number obs.	375,204	122,793	252,411	1,417
Maximum	239.2	239.2	137.5	239.2
95th percentile	28.7	28.6	28.8	78.2
75th percentile	16.3	16.1	16.4	33
Average	12.79	12.72	12.82	24.73
Median	10.7	10.6	10.7	14.4
25th percentile	6.9	6.9	6.9	8.6
5th Percentile	3.7	3.6	3.7	4.9
Minimum	-0.1	-0.1	-0.1	1.9

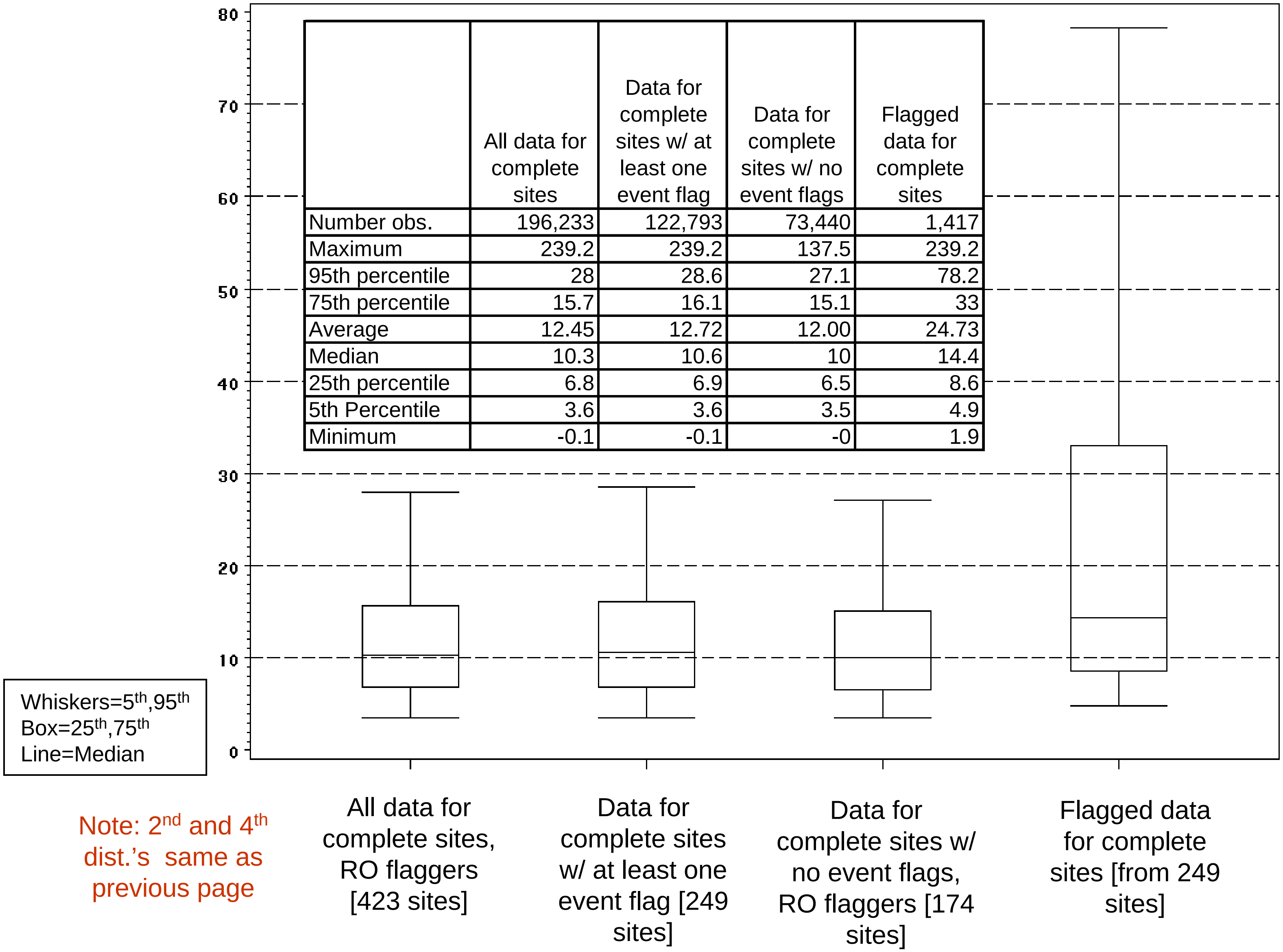


Whiskers=5th,95th
Box=25th,75th
Line=Median

7. Are data distributions similar for sites that flag data vs. sites that don't flag data?

Distribution of PM2.5 concentrations: All data at complete sites for RO flaggers, data for complete sites w/ event flags (RO flaggers), data for complete sites w/out flags (RO flaggers), flagged data from complete sites

	All data for complete sites	Data for complete sites w/ at least one event flag	Data for complete sites w/ no event flags	Flagged data for complete sites
Number obs.	196,233	122,793	73,440	1,417
Maximum	239.2	239.2	137.5	239.2
95th percentile	28	28.6	27.1	78.2
75th percentile	15.7	16.1	15.1	33
Average	12.45	12.72	12.00	24.73
Median	10.3	10.6	10	14.4
25th percentile	6.8	6.9	6.5	8.6
5th Percentile	3.6	3.6	3.5	4.9
Minimum	-0.1	-0.1	-0	1.9



Note: 2nd and 4th dist.'s same as previous page

8. What is specific impact in select areas (case studies)

San Diego (County, MSA, CSA ~ all same)

Site	060730001			060730003			060730006		
Sample freq.	1-3			1-1			1-1		
# of event flags	2			0			2		
# obs.	340			954			330		
% flagged	0.6%			0.0%			0.6%		
	All data	Minus events	Diff.	All data	Minus events	Diff.	All data	Minus events	Diff.
Annual DV	14.6	14.0	0.6	15.7	15.7	0	12.8	12.4	0.4
95th%ile DV	29	28	1	30	30	0	24	24	0
96th%ile DV	30	30	0	32	32	0	26	25	1
97th%ile DV	33	31	2	33	33	0	27	27	0
98th%ile DV	35	34	1	35	35	0	29	28	1
99th%ile DV	38	38	0	38	38	0	31	30	1

Site	060731002			060731007		
Sample freq.	1-1			1-1		
# of event flags	2			4		
# obs.	978			981		
% flagged	0.2%			0.4%		
	All data	Minus events	Diff.	All data	Minus events	Diff.
Annual DV	15.9	15.9	0	15.9	15.6	0.3
95th%ile DV	33	33	0	33	31	2
96th%ile DV	35	34	1	36	34	2
97th%ile DV	36	36	0	38	37	1
98th%ile DV	38	38	0	41	40	1
99th%ile DV	40	40	0	46	45	1

site	value	date	flag
060730001	239.2	10/27/2003	E
060730006	170.2	10/27/2003	E
060731007	170.1	10/27/2003	E
060731007	104.6	10/26/2003	E
060731002	69.2	10/27/2003	E
060731007	42.9	10/29/2003	E
060731007	8.3	10/30/2003	E
060731002	7.3	10/30/2003	E
060730006	5.9	10/30/2003	E
060730001	5.7	10/30/2003	E

•The 3 highest concentrations (in all US) reported by complete sites 2001-2003 were at SD sites and were flagged for forest fires. [Note: 1 site was ‘down’ during the 10/03 episode]

•Removing the flagged data reduces the annual DV’s from 0 to.6ug/m3 (.4 at DV site); percentile DV’s went down 1 to 2 ug/m3 at the high site.

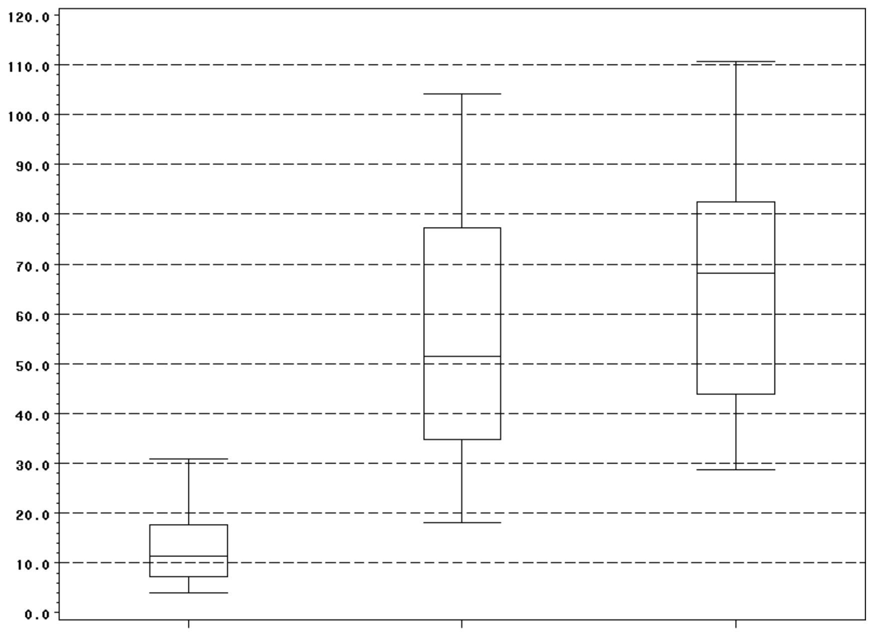
8. What is specific impact in select areas (case studies)

Quebec Fire Event, Effect on PM2.5

- Background:
  - Large smoke plume(s) originating from fires in Quebec, Canada circulated throughout northeast United States in early July 2002 elevating PM2.5 concentrations.
  - General consensus (e.g., CAIR rule modeling) that PM2.5 effects were seen during period, 7/6/2002 through 7/9/2002 in States, NH, VT, CT, RI, MA, NY, NJ, PA, MD, and DE. (Though NC and VA also identified some isolated site-days having effects.) For this analyses, all data during the stated time period and for the stated States were *assumed* to be flagged.
  - Most PM2.5 concentrations for those dates, those states were State flagged (with an 'E' for Forest Fire). Confusion in interpretation of flagging guidance might have prompted some States to not flag. [194 assumed flagged obs, 130 of which were flagged by States]
- Analyses Details:
  - This analyses focuses on the episode effects on 'complete' sites (PM2.5 sites with 11+ samples per Q, all 12 Q's 2001-2003.)
  - 110 complete PM2.5 sites with assumed flags.

8. What is specific impact in select areas (case studies)

Distribution of PM2.5 Concentrations



Concentrations at sites with assumed flags, excluding assumed flagged data (1)

Assumed flagged data (2)

Assumed flagged data also flagged by State (3)

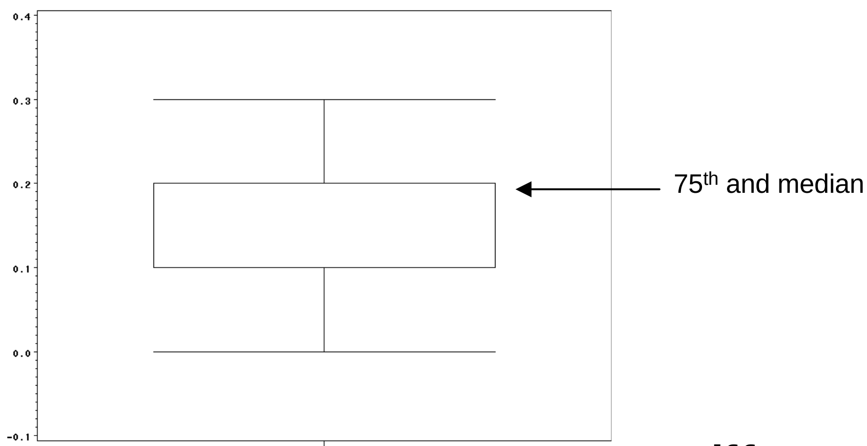
	(1)	(2)	(3)
n	54,716	194	130
MAX	101.7	118.4	118.4
PCT95	30.9	104.1	110.7
PCT75	17.7	77.3	82.5
MEAN	13.7	55.5	66.3
MEDIAN	11.4	51.6	68.3
PCT25	7.3	34.7	43.9
PCT05	4.0	18.1	28.7
MIN	0.0	8.7	11.6

- ‘2’ and ‘3’ distributions *somewhat* similar, hence flag assumption (dates, States) is OK
- 95% of Quebec fire data higher than 75% of other (‘normal’) data. [But some of ‘other’ data flagged for other events!]
- Average Quebec episode (assumed) concentration of 55.5ug/m3; median concentration of 51.6 ug/m3
- The ten highest concentrations reported 2001-2003 for the 110 complete (Quebec impacted) sites were associated with the Quebec event: 104.1 - 118.4 ug/m3.

8. What is specific impact in select areas (case studies)

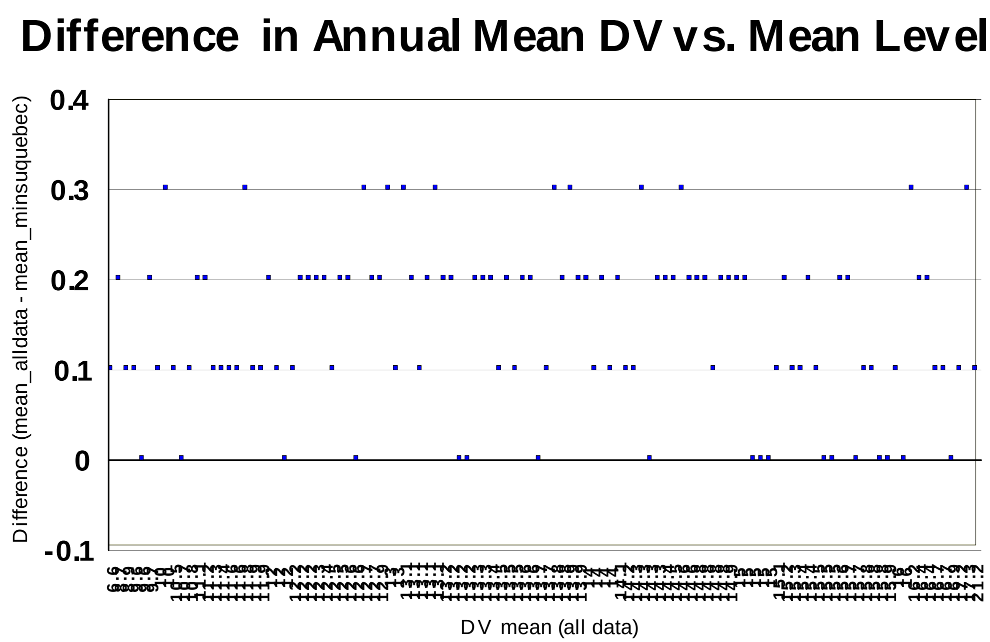
Distribution of Annual Mean DV Differences,

(Mean_alldata minus Mean_minusQuebec)



- If Quebec fire data were excluded, Annual DV's would drop by 0 to .3 ug/m3.
- The average difference is .15 ug/m3
- The median difference is .2 ug/m3
- Sites with higher DV levels do not have bigger differences; see below

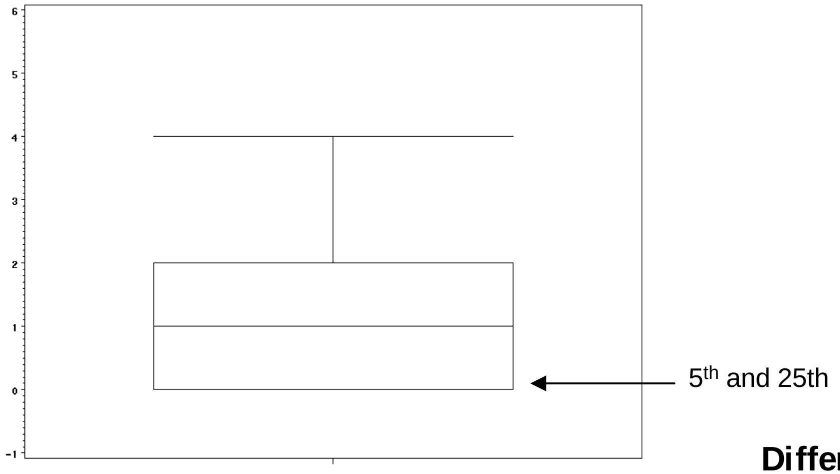
n	110
MAX	0.3
PCT95	0.3
PCT75	0.2
MEAN	0.15
MEDIAN	0.2
PCT25	0.1
PCT05	0
MIN	0



8. What is specific impact in select areas (case studies)

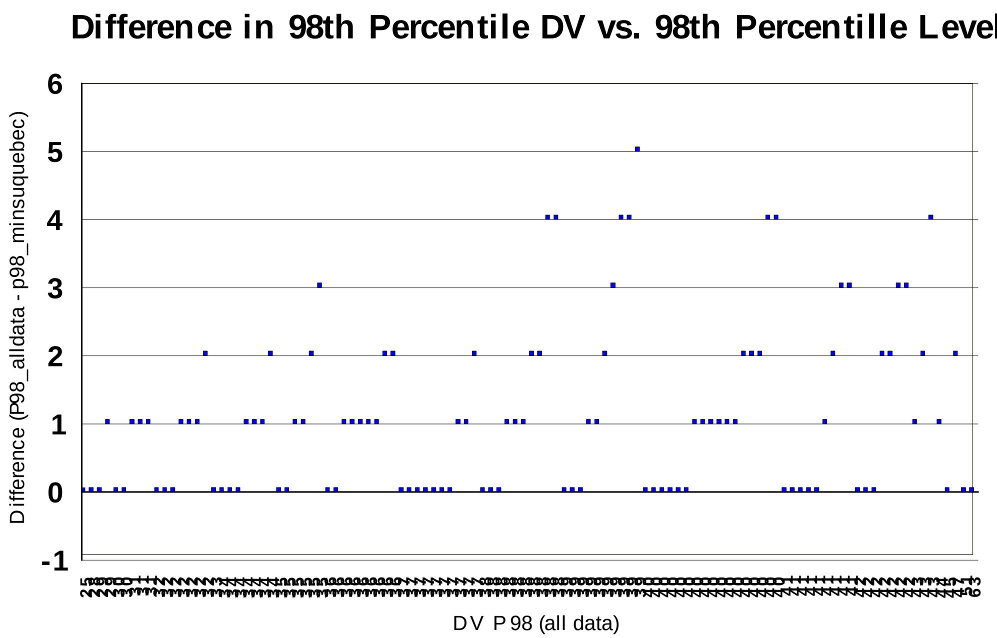
Distribution of 98th Percentile Differences,

(P98_alldata minus P98_minusQuebec)



- If Quebec fire data were excluded, 98th percentile DV's would drop by 0 to 5 ug/m3
- The average difference is 1.07 ug/m3
- The median difference is 1 ug/m3
- Sites with higher DV levels have slightly bigger differences; see below

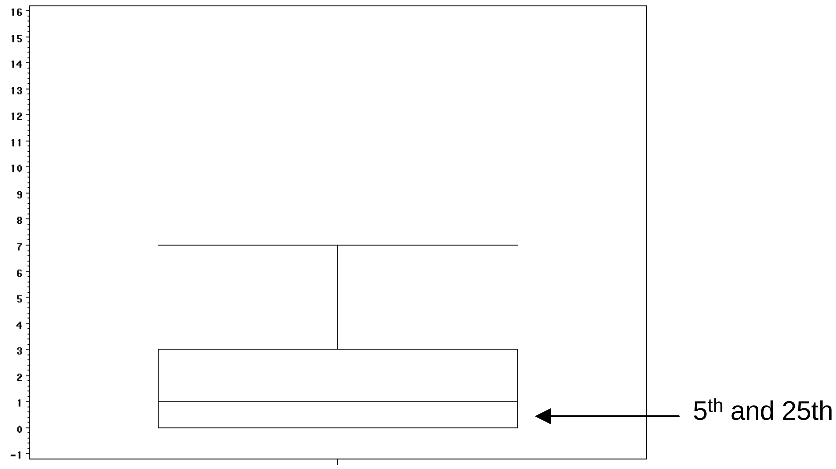
n	110
MAX	5
PCT95	4
PCT75	2
MEAN	1.07
MEDIAN	1
PCT25	0
PCT05	0
MIN	0



8. What is specific impact in select areas (case studies)

Distribution of 99th Percentile Differences,

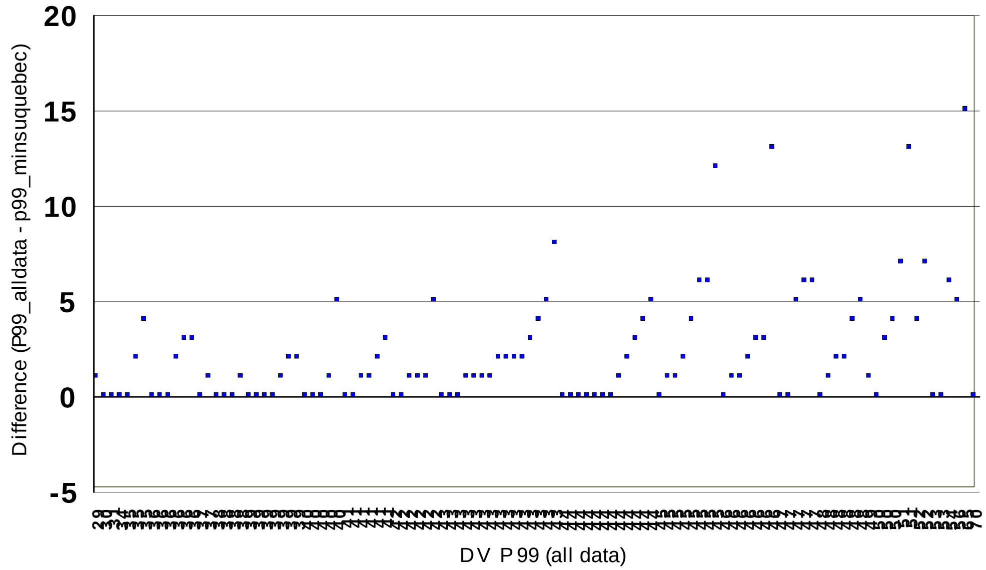
(P99_alldata minus P99_minusQuebec)



- If Quebec fire data were excluded, 99th percentile DV's would drop by 0 to 15 ug/m3
- The average difference is 2.19 ug/m3 (double the avg. diff of 98th)
- The median difference is 1 ug/m3
- Sites with higher DV levels have bigger differences; see below

n	110
MAX	15
PCT95	7
PCT75	3
MEAN	2.19
MEDIAN	1
PCT25	0
PCT05	0
MIN	0

Difference in 99th Percentile DV vs. 99th Percentille Level



8. What is specific impact in select areas (case studies) - Quebec

Crossing Thresholds

Annual Mean DV

	Threshold			
	15	14	13	12
Number of sites with Annual Mean DV (all data) > threshold	26	46	73	87
Number of sites with Annual Mean DV (all data) > threshold but Annual Mean DV (minus Quebec) ≤ threshold	2	2	7	2

98th Percentile DV

	Threshold					
	65	60	55	50	45	40
Number of sites with 98th Percentile DV (all data) > threshold	0	1	1	2	3	24
Number of sites with 98th Percentile DV (all data) > threshold but 98th Percentile DV (minus Quebec) ≤ threshold	0	0	0	0	1	9

99th Percentile DV

	Threshold					
	65	60	55	50	45	40
Number of sites with 99th Percentile DV (all data) > threshold	1	2	3	10	32	79
Number of sites with 99th Percentile DV (all data) > threshold but 99th Percentile DV (minus Quebec) ≤ threshold	0	1	2	6	14	16

- 98th percentile more stable of indicator than 99th percentile

Episodic Events – PM10-2.5

- Questions:
  1. What types of events are flagged in AQS?
  2. Are there a significant amount of event-flagged data in AQS?
  3. Do ‘high’ sites flag more data than ‘low’ sites?
  4. How do events impact DV’s
  5. Are the impacts different for ‘high’ vs ‘low’ sites?
  6. Are data distributions similar for sites that flag data vs. sites that don’t flag data?

1. What types of events are flagged in AQS?

PM10-2.5 Flag Counts, 2001-2003 - Data for sites that meet completeness

Flag Description	Event Class	Flag Count	Percent of Event Flags	Number of Sites Reporting
Forest Fire	natural	228	24.3%	102
Highway Construction	except.	208	22.2%	3
Sahara Dust	natural	189	20.1%	6
Construction/Demolition	except.	107	11.4%	14
High Winds	natural	90	9.6%	34
Volcanic Eruptions	natural	60	6.4%	6
Roofing Operations	except.	15	1.6%	3
Infrequent Large Gatherings	except.	14	1.5%	7
Rerouting Of Traffic	except.	13	1.4%	1
Agricultural Tilling	except.	7	0.7%	3
Clean Up After A Major Disaster	except.	4	0.4%	4
Sandblasting	except.	1	0.1%	1
Prescribed Burning	except.	1	0.1%	1
Seismic Activity	natural	1	0.1%	1
Total		938		146

Notes:
- **Complete sites defined as those with 12, 8, or 4 consecutive quarters of 11+ samples. [489 sites total]**
- **PM10-2.5 flag was set to the PM10 flag if it exists, else to the PM2.5 flag if it exists.**

- Most flagged data relate to natural events (~ 60%)
- Forest fires is the most common event flagged (looking by flag) but ‘highway construction, ‘Sahara dust’, ‘Construction/Demolition’, and ‘High Winds’ area also common (over 100+ site days)
- Looking by site, ‘Forest Fires’ affected the most sites, followed by ‘High Winds’
- 30% of all sites (146 / 489) reported at least one event flag.
- 2 of the 10 highest PMc values were flagged; 37 of the top 100; and 99 of the top 1000
- All ‘Volcanic Eruptions’ and ‘Sahara Dust’ flags are associated with monitoring site in Puerto Rico and the Virgin Islands

2. Are there a significant amount of event-flagged data in AQS?

Percentage of event-flagged data - for complete sites [489]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	0.9%	3.8%	4.2%	4.8%	5.3%	7.2%

All complete sites…. But some sites may not flag. (Hence, perhaps biased low). Reference as ‘complete sites’

Percentage of event-flagged data - at sites with at least 1 flagged point [146 sites]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	3.1%	12.8%	14.0%	16.0%	17.9%	24.1%

Complete sites with at least one event flag 2001-2003. But some other sites may have flagged if event occurred. (Hence, perhaps biased high). Reference as ‘flag sites’

Percentage of event-flagged data at complete sites where RO has at least 1 flagged datapoint (not necessarily at all sites) [327 sites]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	1.4%	5.7%	6.2%	7.2%	8.0%	10.8%

Complete sites where the reporting organization (RO) has flagged at least one event 2001-2003, though not necessarily at all sites. Reference as ‘RO flaggers’. [The RO knows how to flag.]

- Event-flagged data account for .9% to 3.1% of all reported observations. (about triple the PM2.5 rate)
- They account for considerably higher percentage of high values (i.e., values $\geq$ 95th, 96th, 97th, 98th, 99th percentile)

3. Do ‘high’ sites ($\geq$ 50 98th percentile*) flag more data than ‘low’ sites (< 50 98th percentile*)?

Percentage of event-flagged data at complete sites where RO has at least 1 flagged datapoint (not necessarily at all sites) [327 sites]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	1.4%	5.7%	6.2%	7.2%	8.0%	10.8%

Same as previous page (bottom), ‘RO flaggers’. Break out by high / low.

Percentage of event-flagged data - RO flaggers, sites $\geq$ 50 [68 sites]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	3.7%	14.9%	16.5%	19.4%	21.4%	29.6%

High sites

Percentage of event-flagged data - RO flaggers, sites < 50 [259 sites]

	Percent of All Samples	Percent of Samples Equal or Above Percentile Metric				
		95th	96th	97th	98th	99th
Site Average	0.7%	3.3%	3.5%	3.9%	4.5%	5.9%

Low sites

- High sites flag more data. (The flagged data makes them ‘high’ sites.). They flag about 5 times in total (on average) and also, 5 times the number of extreme values

** Approximately 20% of the 489 sites in the 2001-2003 PM10-2.5 database have a 98th percentile $\geq$ 50.*

4. How do events impact DV's?

Reductions (µg/m3) in annual and 24-hour design values as a result of exempting event-flagged data - complete sites [489 sites]

	Reduction (ug/m3) in Stated Metric					
Site change	Annual DV	95th Percentile	96th Percentile	97th Percentile	98th Percentile	99th Percentile
Maximum	10.8	58	137	215	202	189
95th percentile	1.3	3	3	5	8	21
75th Percentile	0.11	0	0	0	0	0
Average	0.17	0.68	0.97	1.41	1.65	3.30
Median	0.01	0	0	0	0	0
25th percentile	-0.06	0	0	0	0	0
5th Percentile	-0.29	0	0	0	0	0
Minimum	-1.39	0	0	0	0	0

Reductions (µg/m3) in annual and 24-hour design values as a result of exempting event-flagged data - flag sites [146 sites]

	Reduction (ug/m3) in Stated Metric					
Site change	Annual DV	95th Percentile	96th Percentile	97th Percentile	98th Percentile	99th Percentile
Maximum	10.76	58	137	215	202	189
95th percentile	2.33	6	11	20	23	80
75th Percentile	0.51	2	2	2	3	8
Average	0.52	2.29	3.25	4.72	5.53	11.05
Median	0.09	0	0	0	0	0
25th percentile	-0.02	0	0	0	0	0
5th Percentile	-0.23	0	0	0	0	0
Minimum	-1.39	0	0	0	0	0

Reductions (µg/m3) in annual and 24-hour design values as a result of exempting event-flagged data - RO flaggers [327 sites]

	Reduction (ug/m3) in Stated Metric					
Site change	Annual DV	95th Percentile	96th Percentile	97th Percentile	98th Percentile	99th Percentile
Maximum	10.8	58	137	215	202	189
95th percentile	1.6	5	5	10	12	26
75th Percentile	0.14	0	0	0	0	0
Average	0.25	1.02	1.45	2.11	2.47	4.93
Median	0.02	0	0	0	0	0
25th percentile	-0.05	0	0	0	0	0
5th Percentile	-0.28	0	0	0	0	0
Minimum	-1.39	0	0	0	0	0

- The last table (RO flaggers) probably represents the best guess at national average effect.
- On average, removing flagged data would reduce annual DV's by about .25 ug/m3, 98^{th} percentiles by about 2-3 ug/m3, and 99^{th} percentiles by 4-5 ug/m3.
- Some sites would have very large changes in in percentiles (95^{th}-99^{th}) if flagged data were omitted; see max and 95^{th}%ile site change rows

5. Are the impacts different for ‘high’ vs ‘low’ sites?

Reductions (µg/m3) in annual and 24-hour design values as a result of exempting event-flagged data - RO flaggers, sites ≥ 50 [68 sites]

	Reduction (ug/m3) in Stated Metric					
Site change	Annual DV	95th Percentile	96th Percentile	97th Percentile	98th Percentile	99th Percentile
Maximum	10.8	58	137	215	202	189
95th percentile						
75th percentile						
Average	0.88	3.82	5.94	8.76	10.28	19.65
Median	0.24	0.5	0	0.5	0	0
25th percentile						
5th Percentile						
Minimum	-1.39	0	0	0	0	0

Reductions (µg/m3) in annual and 24-hour design values as a result of exempting event-flagged data - RO flaggers, sites < 50 [259 sites]

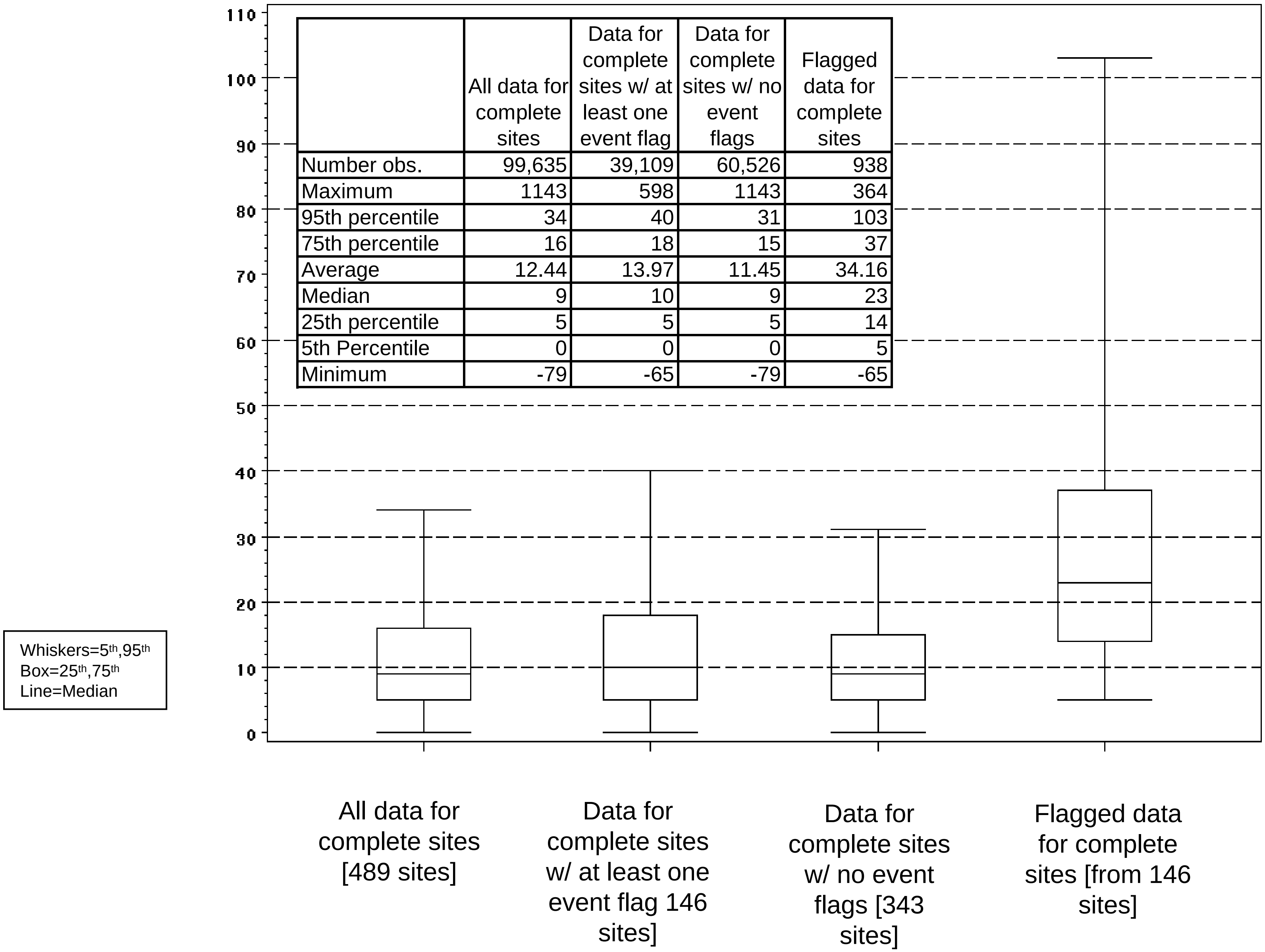
	Reduction (ug/m3) in Stated Metric					
Site change	Annual DV	95th Percentile	96th Percentile	97th Percentile	98th Percentile	99th Percentile
Maximum	3.35	24	24	24	24	37
95th percentile						
75th percentile						
Average	0.09	0.29	0.27	0.36	0.42	1.07
Median	0.01	0	0	0	0	0
25th percentile						
5th Percentile						
Minimum	-1.11	0	0	0	0	0

•Apparent differences in effect on annual DV and percentile DV’s
•High sites have about ten times the reduction in annual DV’s… about .9ug/m3 on average
•High sites have 10-20 times the reduction in percentile DV’s
•Some sites (high and low) have considerable effects

6. Are data distributions similar for sites that flag data vs. sites that don't flag data?

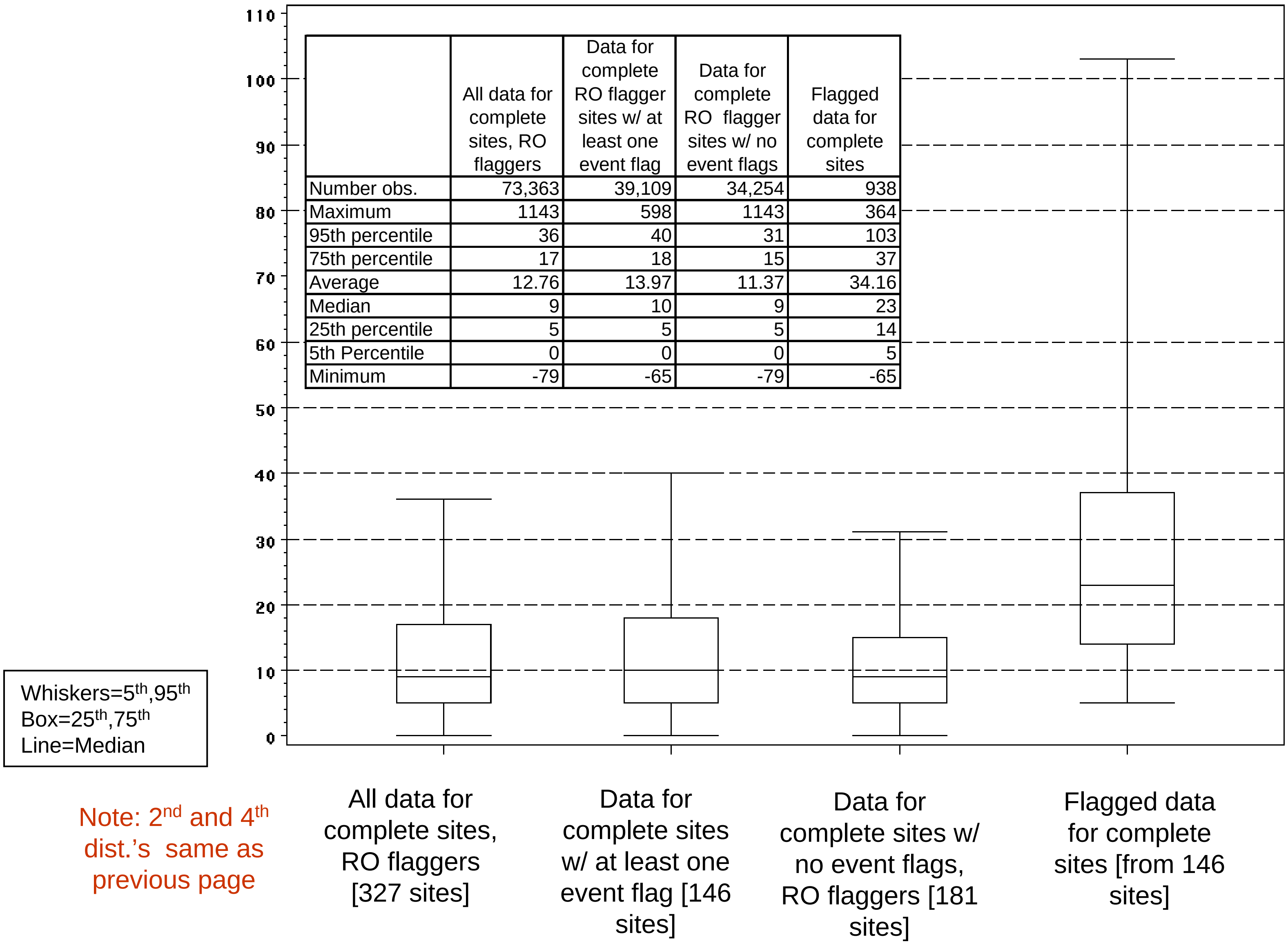
- See next 2 slides
- 2nd slide more accurate comparison (RO flaggers)
- Some differences on high end of distributions (flag sites vs. no flag sites). Flag site data higher.
- Obvious differences in data distributions of all data vs. flagged data.
    - Flagged data generally higher, average concentration is 12.4 - 12.8 for all data (at comp sites) vs. 34.1 for flagged data

7. Distribution of PM10-2.5 concentrations: All data at complete sites, data for complete sites w/ event flags, data for complete sites w/out flags, flagged data from complete sites



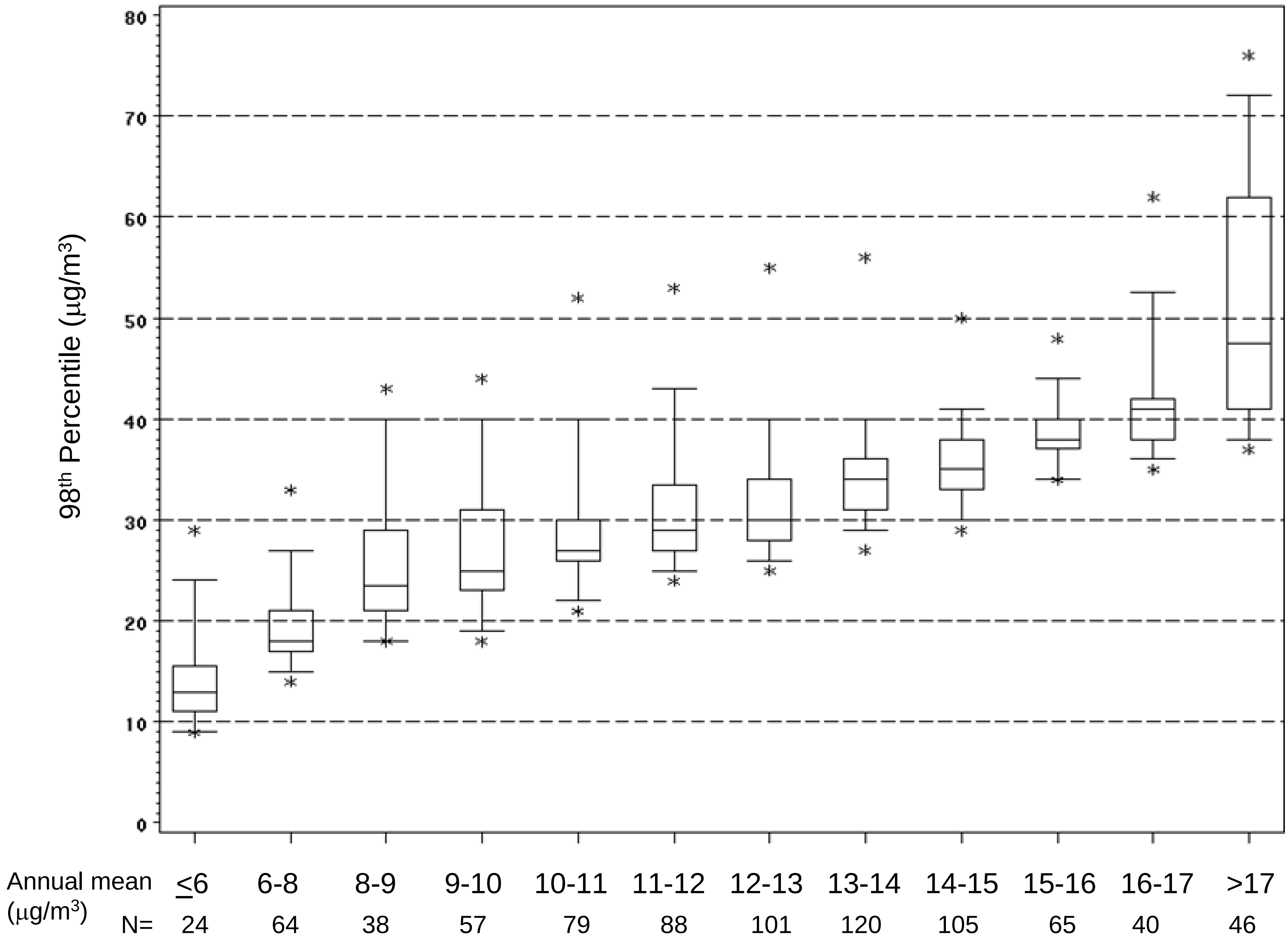
	All data for complete sites	Data for complete sites w/ at least one event flag	Data for complete sites w/ no event flags	Flagged data for complete sites
Number obs.	99,635	39,109	60,526	938
Maximum	1143	598	1143	364
95th percentile	34	40	31	103
75th percentile	16	18	15	37
Average	12.44	13.97	11.45	34.16
Median	9	10	9	23
25th percentile	5	5	5	14
5th Percentile	0	0	0	5
Minimum	-79	-65	-79	-65

7. Distribution of PM10-2.5 concentrations: All data at complete sites for RO flaggers, data for complete sites w/ event flags (RO flaggers), data for complete sites w/out flags (RO flaggers), flagged data from complete sites



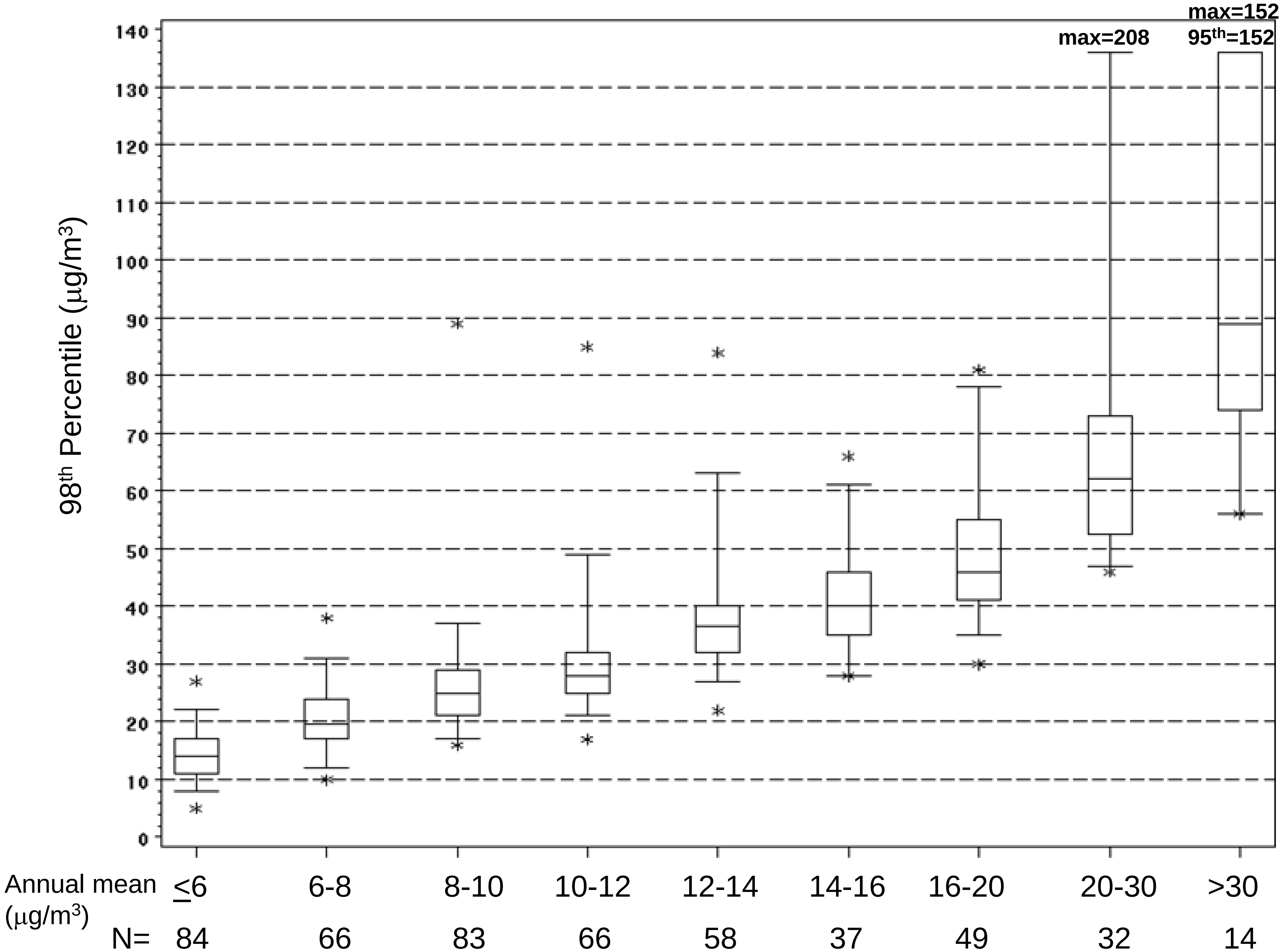
	All data for complete sites, RO flaggers	Data for complete RO flagger sites w/ at least one event flag	Data for complete RO flagger sites w/ no event flags	Flagged data for complete sites
Number obs.	73,363	39,109	34,254	938
Maximum	1143	598	1143	364
95th percentile	36	40	31	103
75th percentile	17	18	15	37
Average	12.76	13.97	11.37	34.16
Median	9	10	9	23
25th percentile	5	5	5	14
5th Percentile	0	0	0	5
Minimum	-79	-65	-79	-65

Note: 2nd and 4th dist.'s same as previous page



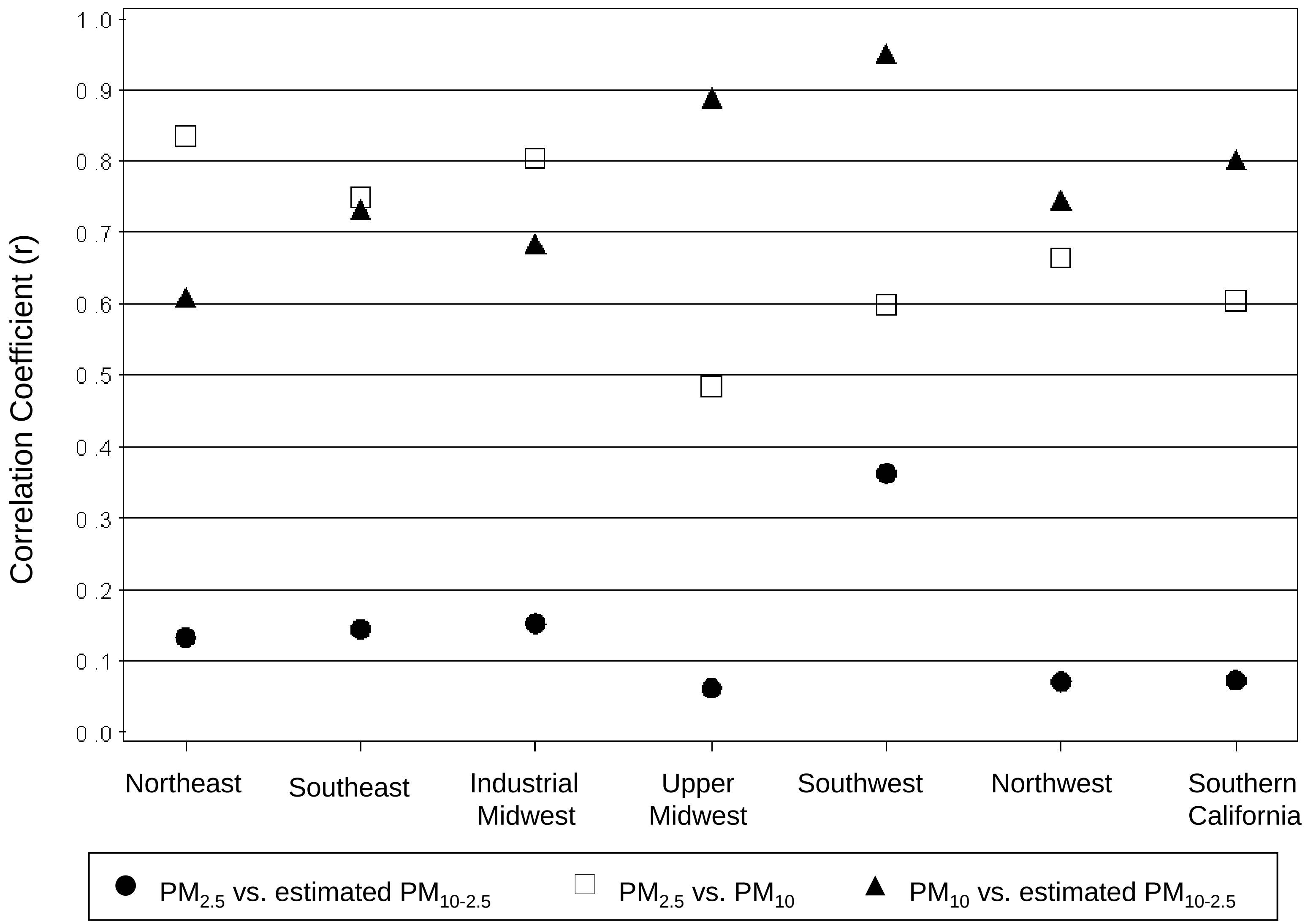
Distribution of annual mean vs. 98th percentile 24-hour average $PM_{2.5}$ concentrations, 2001-2003.

Box depicts interquartile range and median; whiskers depict 5th and 95th percentiles; asterisks depict minima and maxima. N= number of sites.

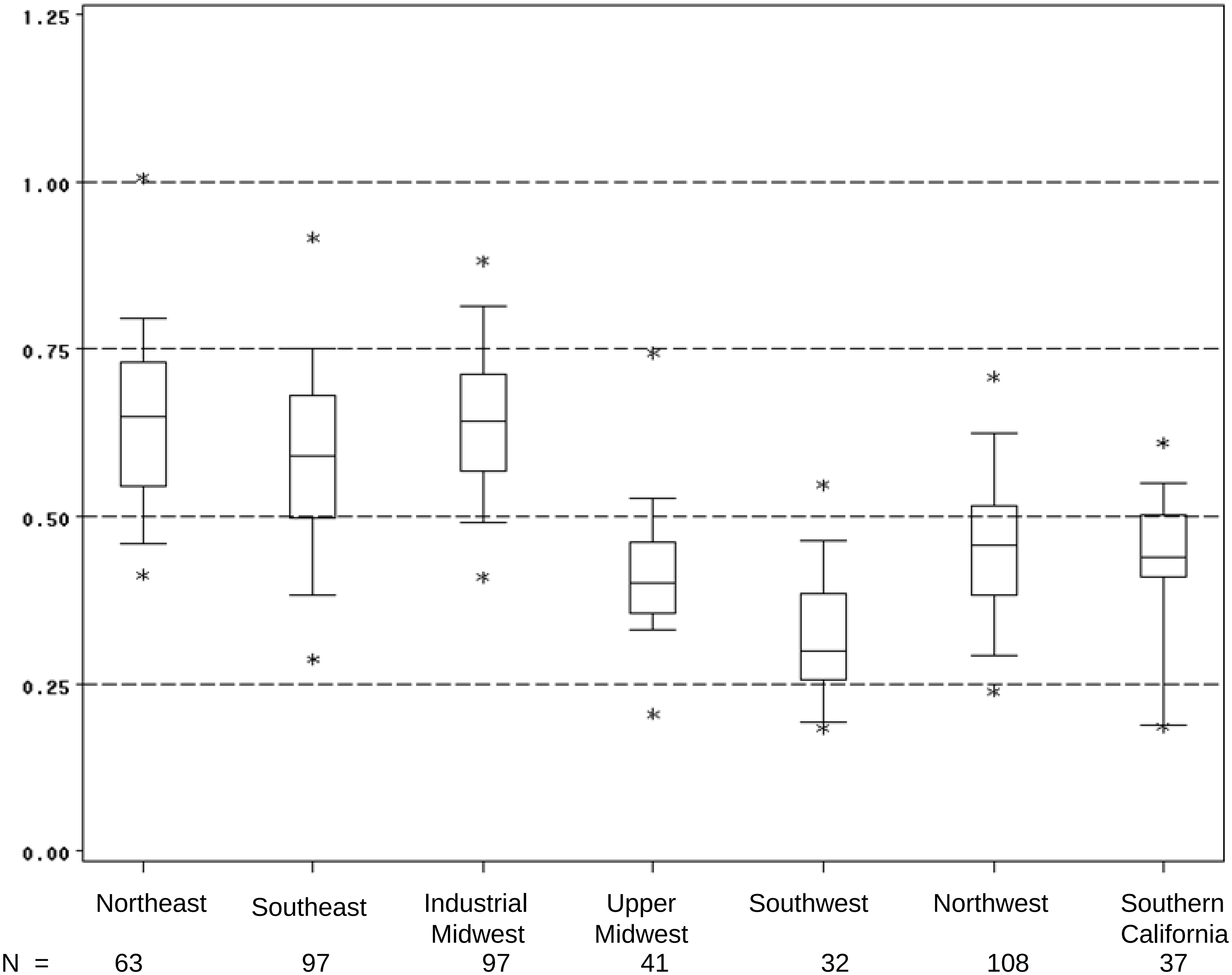


Distribution of estimated annual mean vs. 98th percentile 24-hour average $PM_{10\text{-}2.5}$ concentrations, 200- 2003.

Box depicts interquartile range and median; whiskers depict 5th and 95th percentiles; asterisks depict minima and maxima. N= number of sites.



Regional average correlations of 24-hour average PM by size fraction.



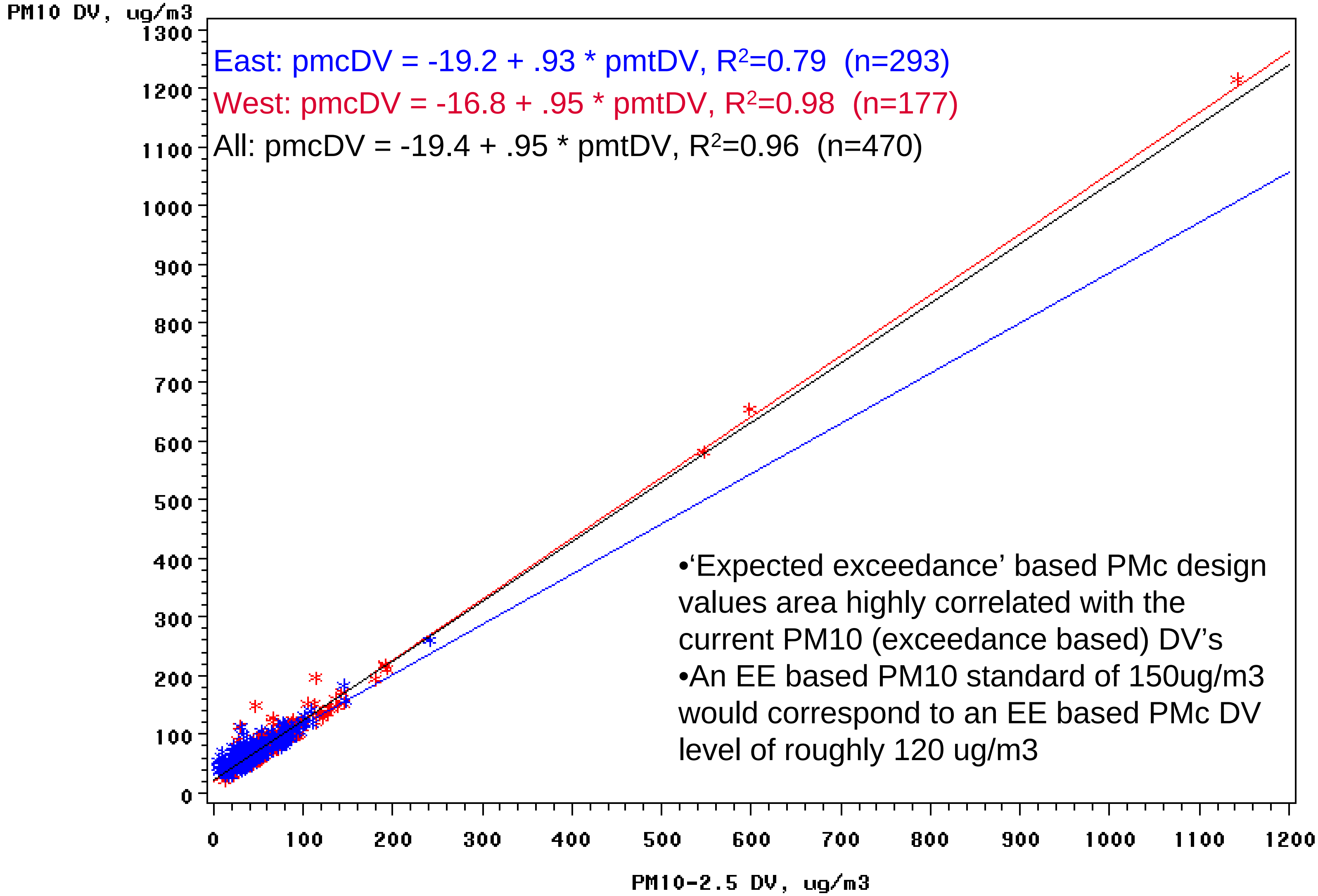
Distribution of ratios of $PM_{2.5}$ to PM_{10} by region, 2001-2003.

Box depicts interquartile range and median; whiskers depict 5th and 95th percentiles; asterisks depict minima and maxima. N = number of sites.

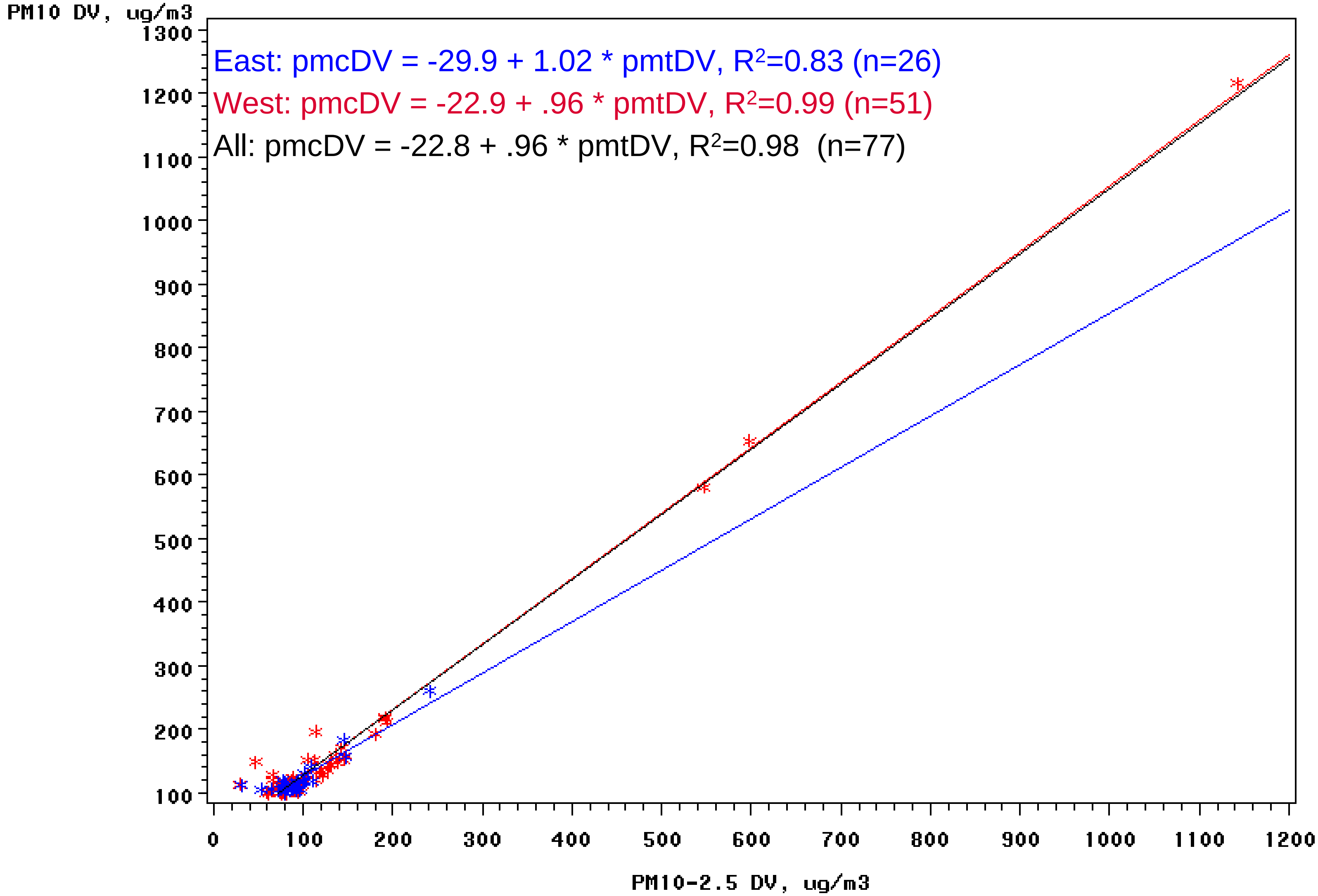
PM10-2.5 Equivalence to PM10 NAAQS (Daily Standard)

- Questions:
  1. How equivalent (well correlated) are various PMc design value type metrics to the existing PM10 24-hour standard?
  2. Are the relationships different for high PM10 levels?
  3. Do the relationships vary across regions?
  4. What levels (for the various PMc design value type metrics) would correspond to the 150 ug/m3 level for the current PM10 24-hour standard.
- Analyses details:
  1. Comparative 'design value' type statistics for PM10 and PM10-2.5 were all constructed from the 489 site 2001-2003 PMc database.
     - This database required 12, 8, or 4 consecutive quarters of 11+ samples. Only the data from the 12,8,4 quarter period were utilized. The percentile based DV's were calculated by averaging the associated annualized (4-quarter increment) statistics. [E.g., Assume a site had 8 complete quarters starting with 2001-Q3 and ending with 2003-Q2. Two 'annual' 98^{th} percentiles were computed, one for 2001-Q3 through 2002-Q2 and the other for 2002-Q3 through 2003-Q2. These 2 'annual' numbers were then averaged to obtain the site's 98^{th} percentile design value.
     - Actual 2001-2003 PM10 24-hour design values were also obtained and matched to the 489 sites. These numbers, termed 'real' or 'actual', are used in some of the analyses for perspective only.
     - In addition to the possible mismatch with time period, the 'actual' and comparative PM10 expected exceedance DV's might also not match due to the following processing differences. 1) Actual DV's are computed at monitor level, comparative DV'S were calculated at site level (averaging same site, same day FRM measurements); and 2) actual DV's exclude Regionally-concurred event flags but comparative DV's exclude all event flags, concurred or not.

PM10 DV (Expected Exceedance form) versus PM10-2.5 DV (EE form) all pmt levels

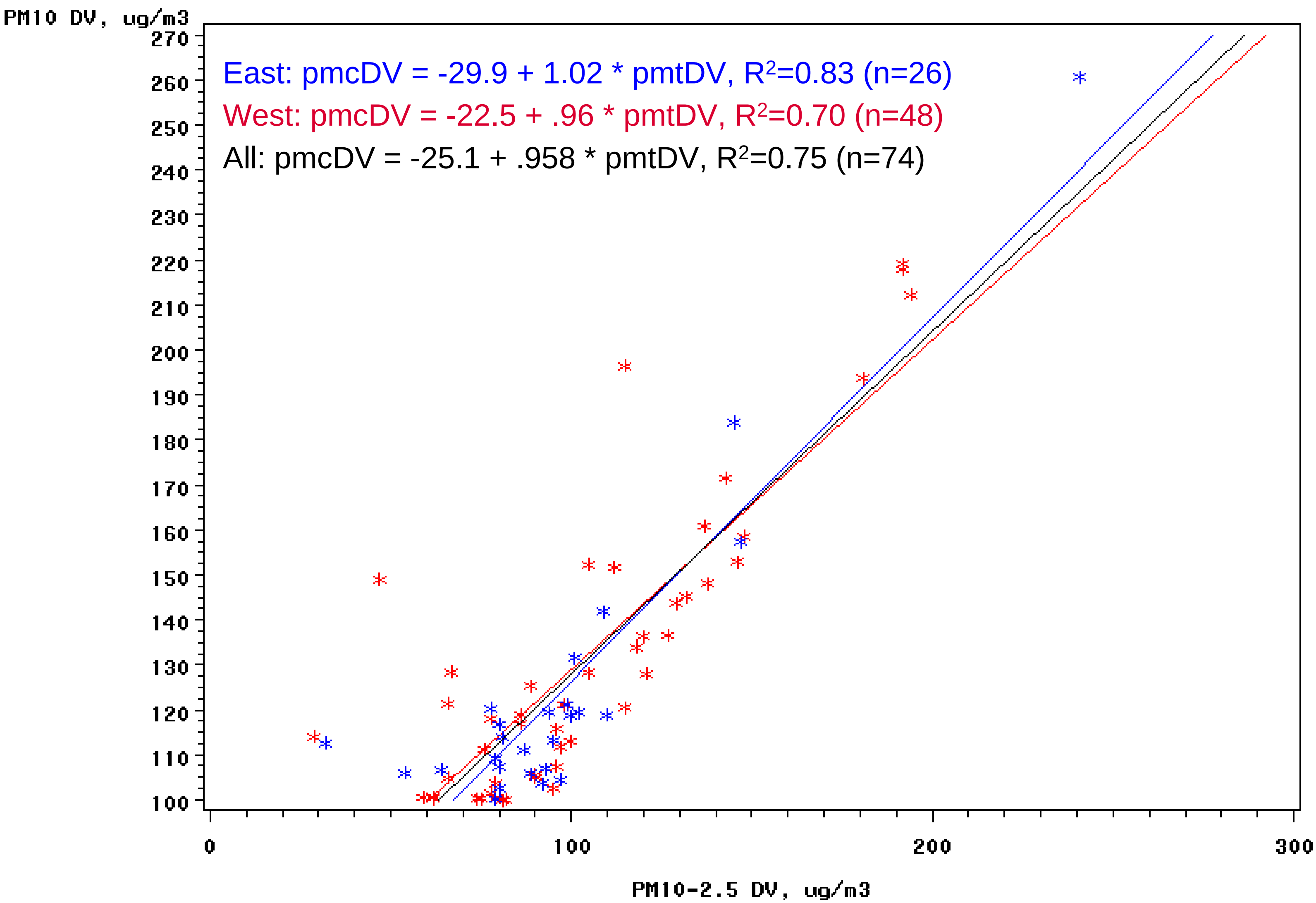


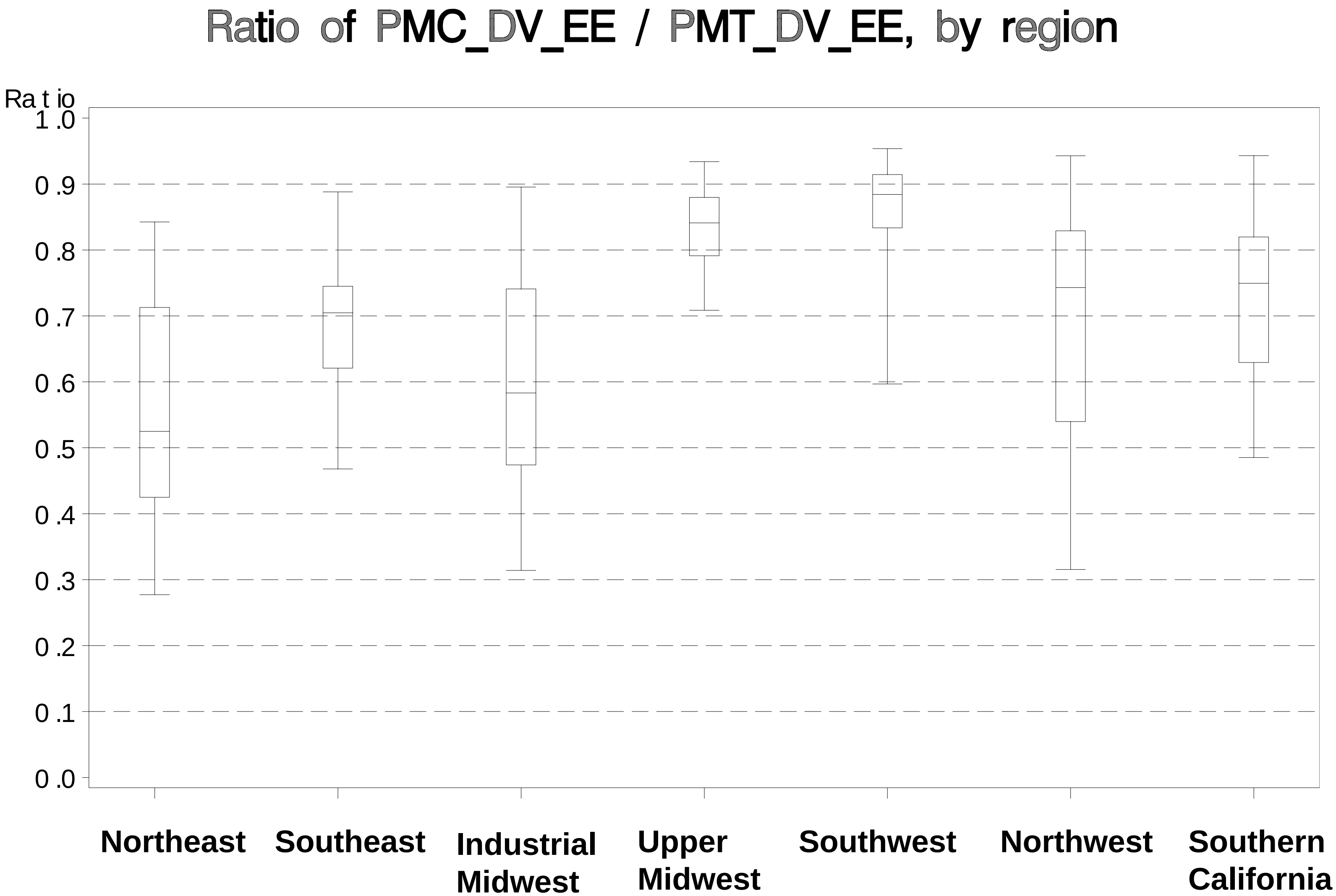
PM10 DV (Expected Exceedance form) versus PM10-2.5 DV (also EE form)
pmt ge 100



PM10 DV (Expected Exceedance form) versus PM10-2.5 DV (also EE form)

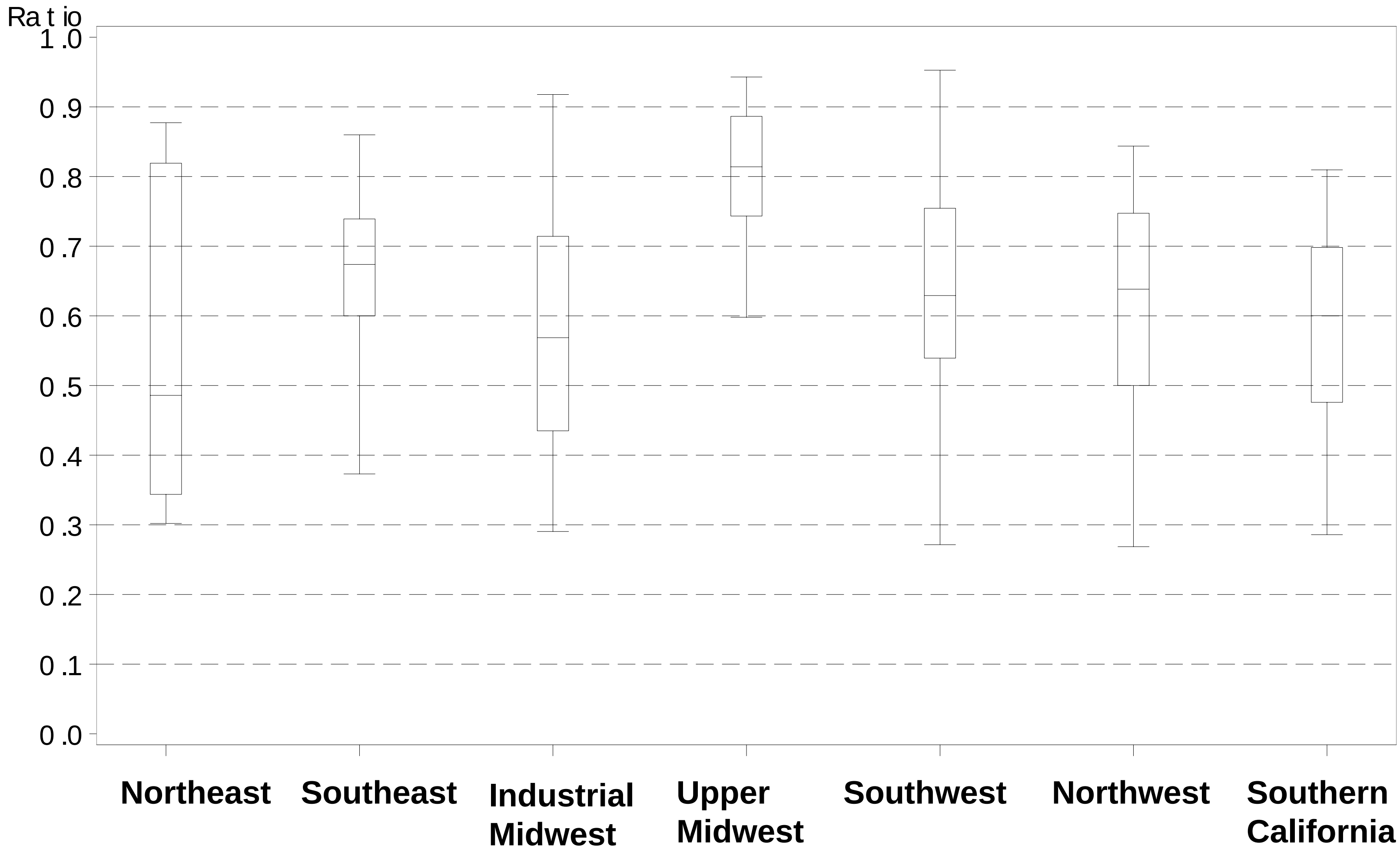
pmt ge 100, le 300



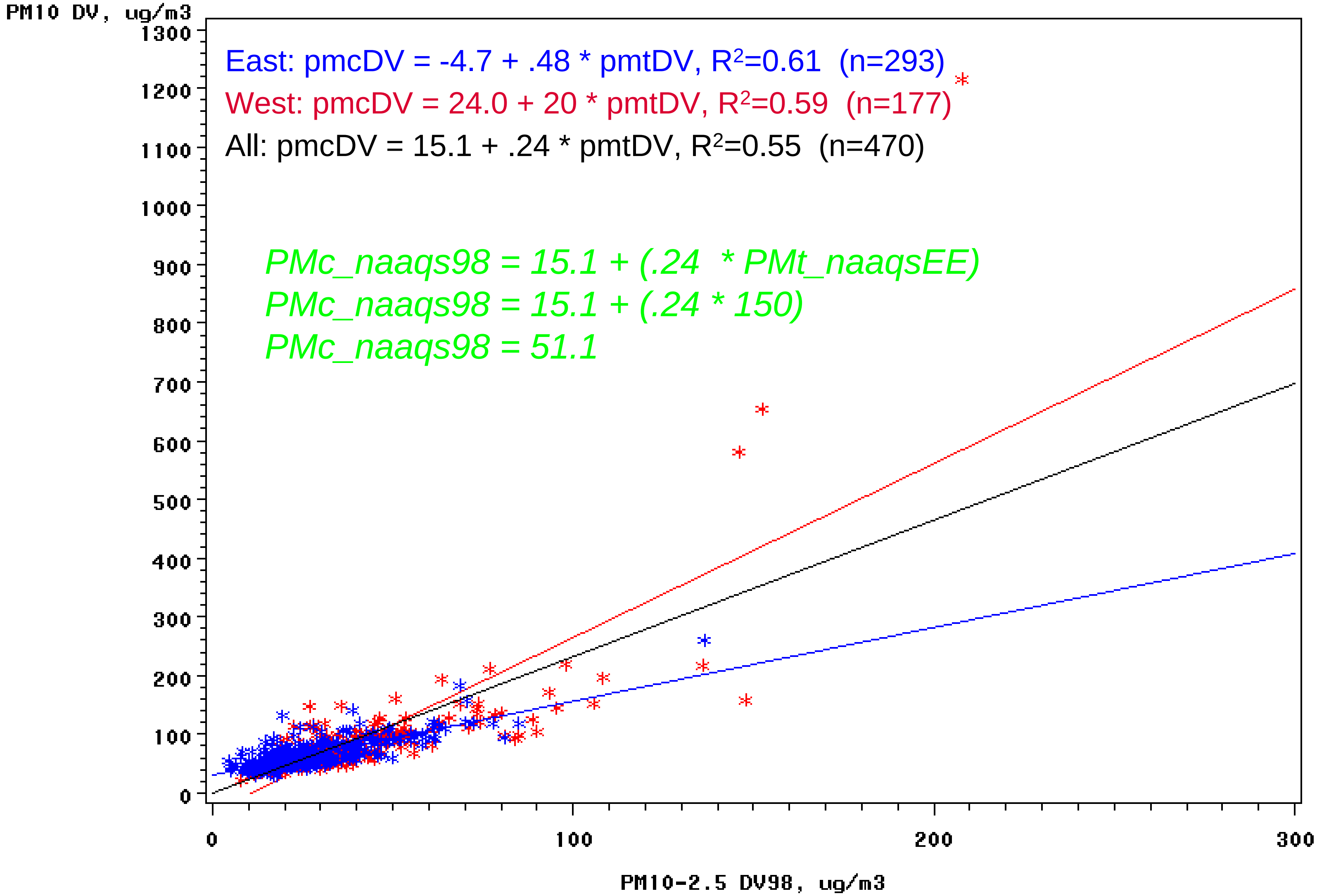

Ratio of PMC_DV_EE / PMT_DV_EE, by region
Ratio
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Northeast
Southeast
Industrial Midwest
Upper Midwest
Southwest
Northwest
Southern California

For perspective! – Similar to previous page

Ratio of PMC_DV_EE (12,8,4) / PMT_DV_EE (real0103), by region

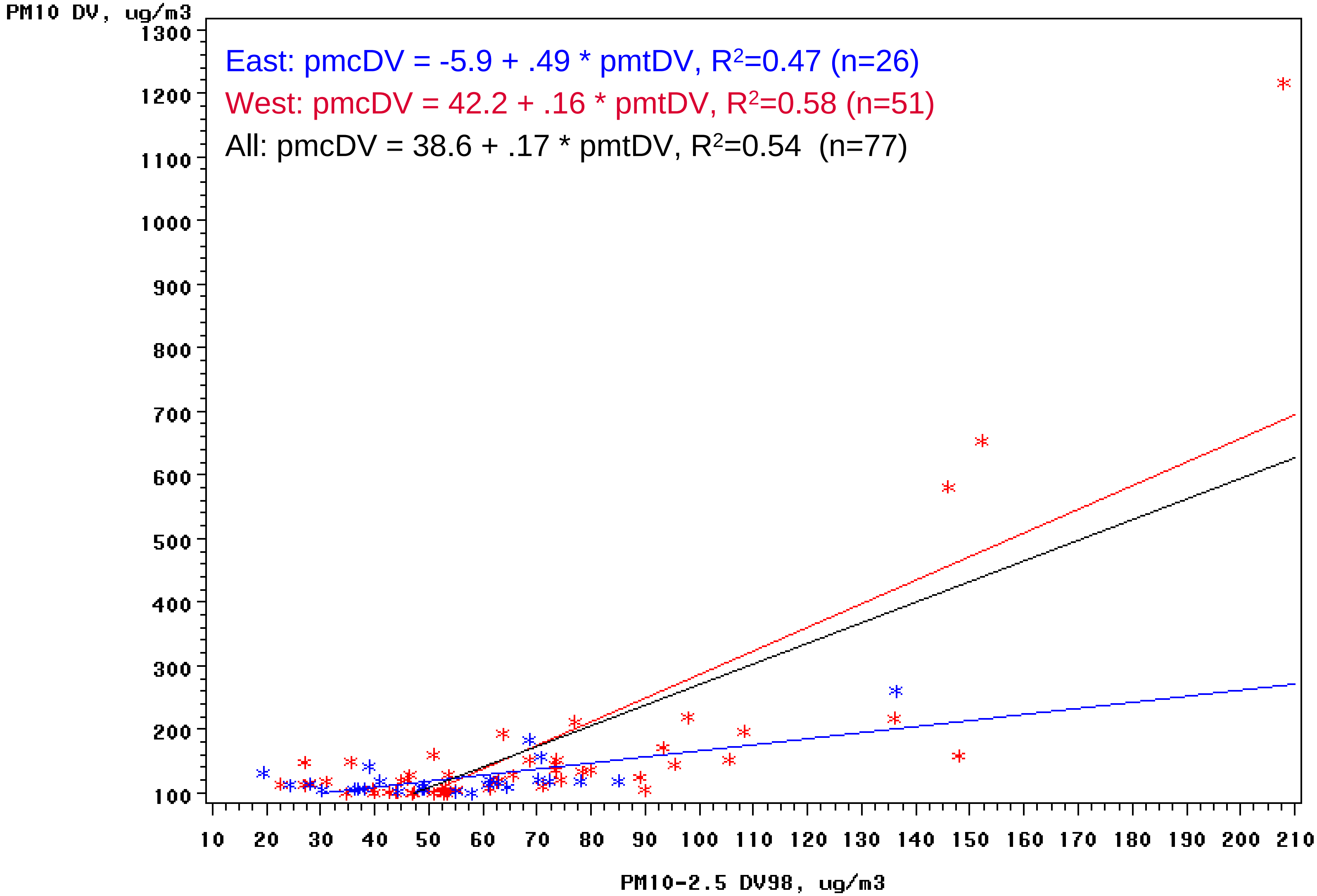


PM10 DV (EE form) versus PM10-2.5 DV (98th percentile form)
all pmt levels

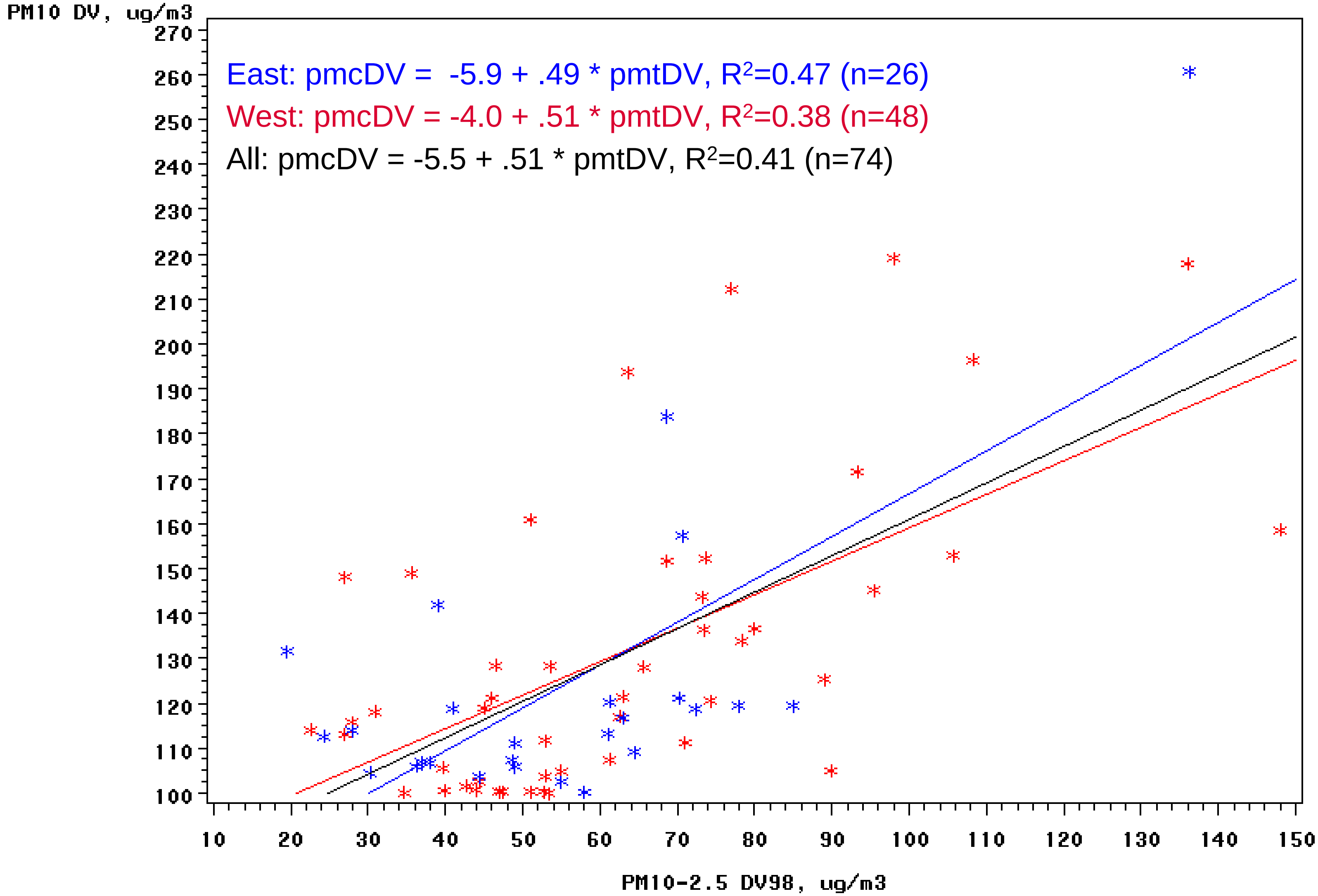


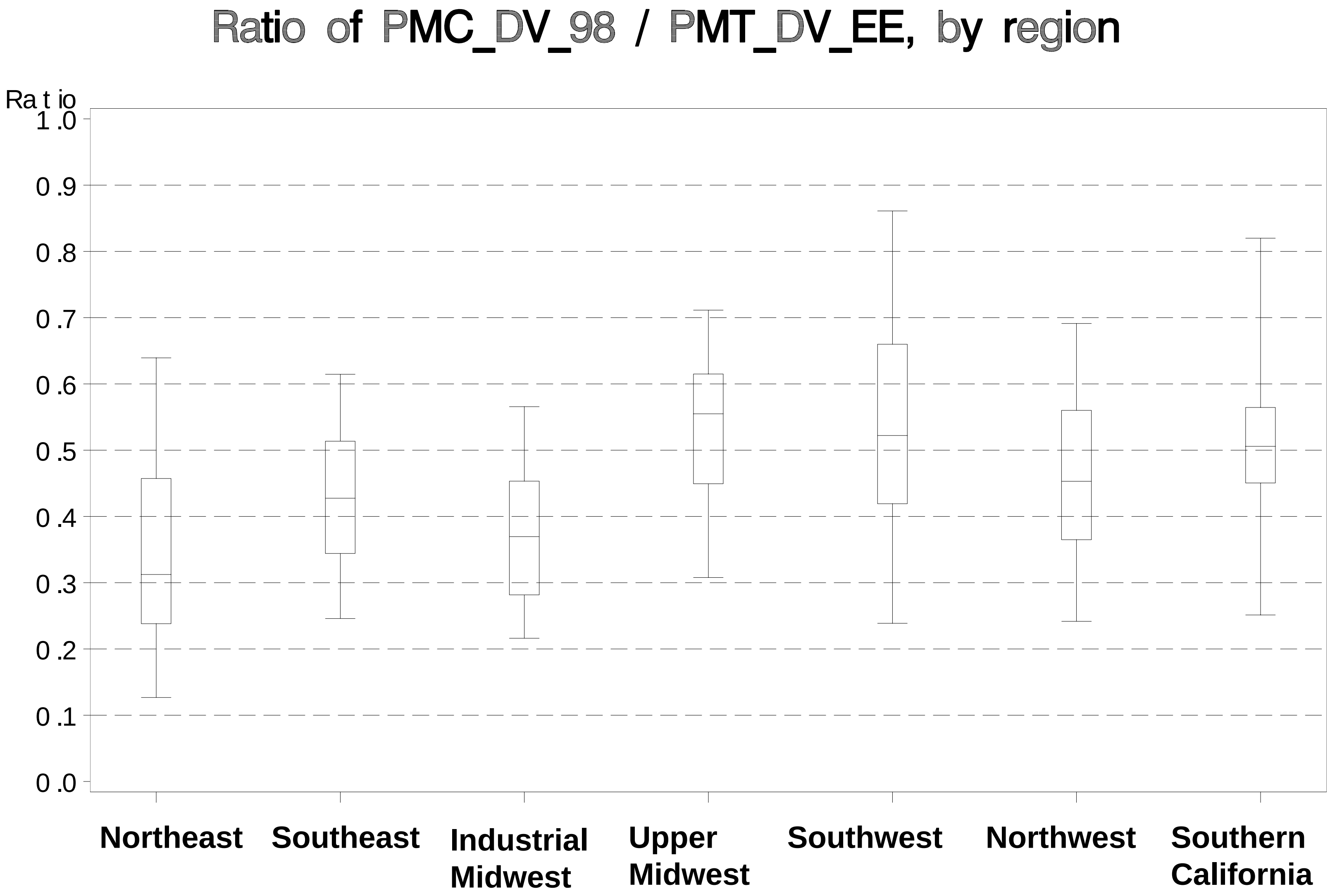
PM10 DV (EE form) versus PM10-2.5 DV (98th percentile form)

pmt ge 100



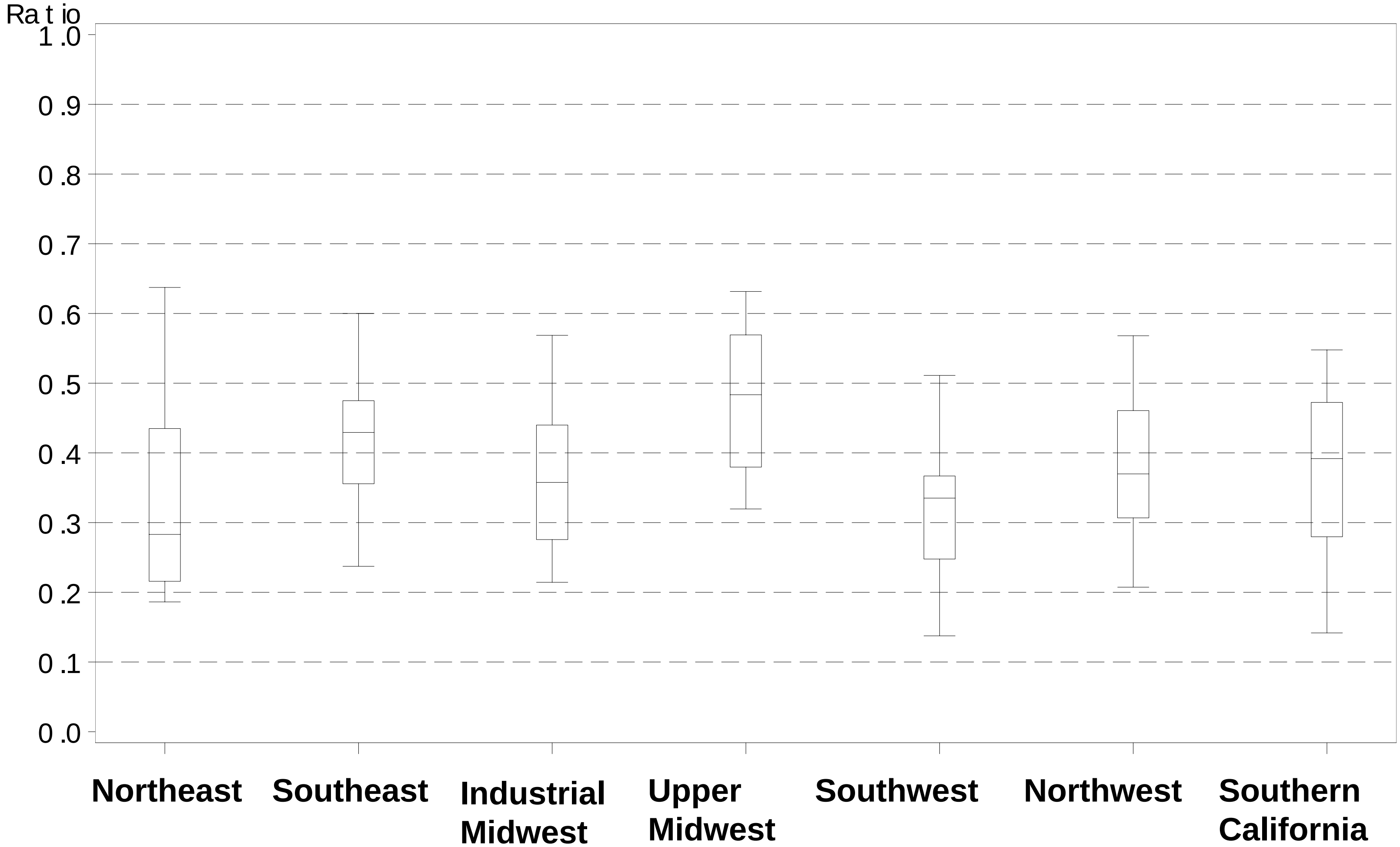
PM10 DV (EE form) versus PM10-2.5 DV (98th percentile form)
pmt ge 100, le 300



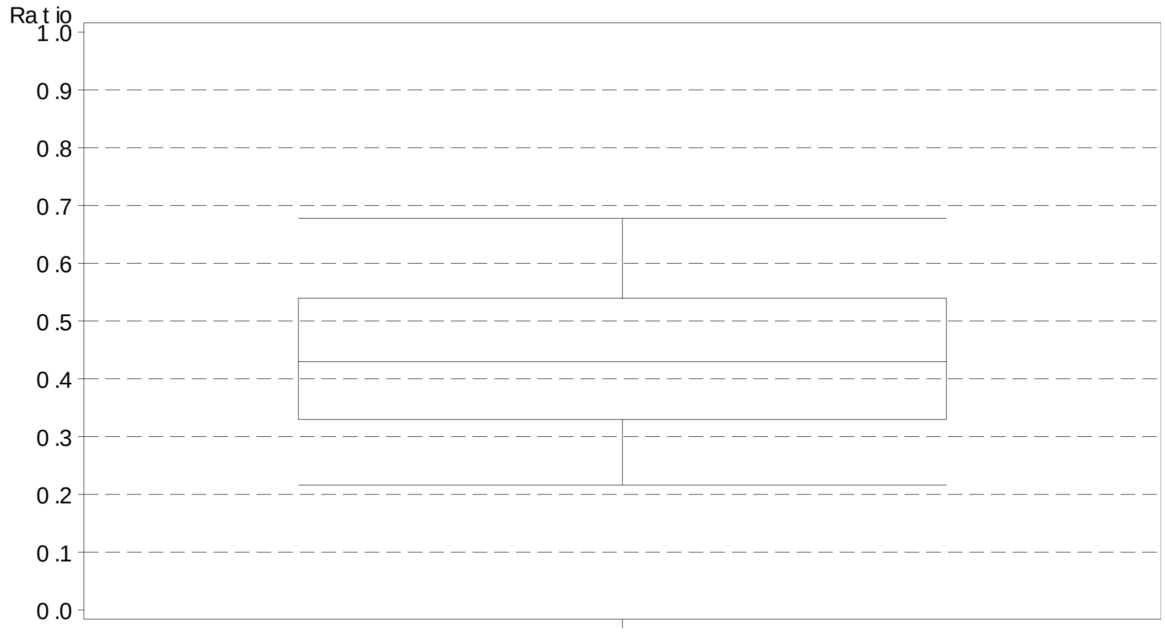

Ratio of PMC_DV_98 / PMT_DV_EE, by region
Ratio
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Northeast
Southeast
Industrial Midwest
Upper Midwest
Southwest
Northwest
Southern California

For perspective! – Similar to previous page

Ratio of PMC_DV_98 (12,8,4) / PMT_DV_EE (real0103), by region

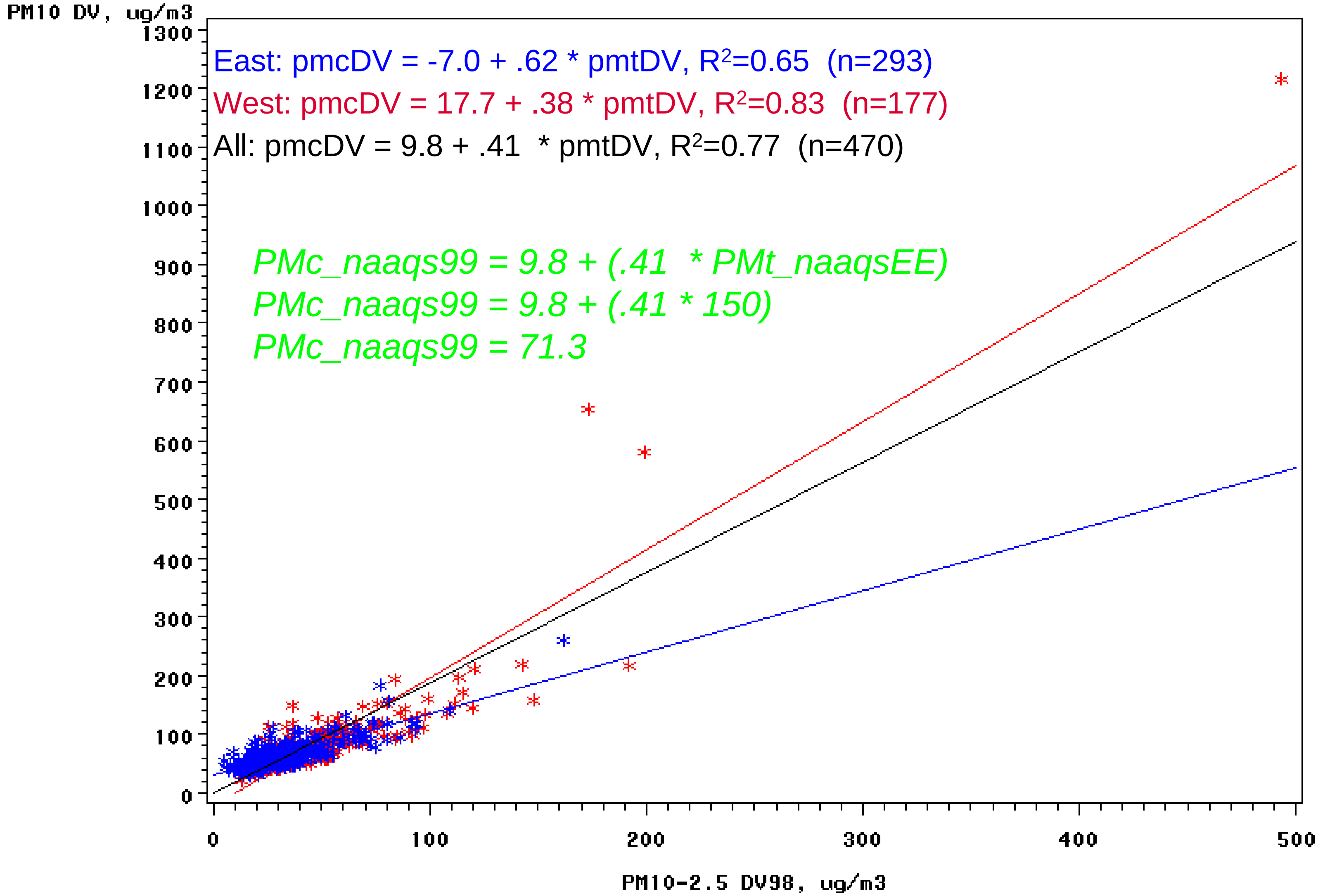


Ratio of PMC_DV_98 / PMT_DV_EE, all sites

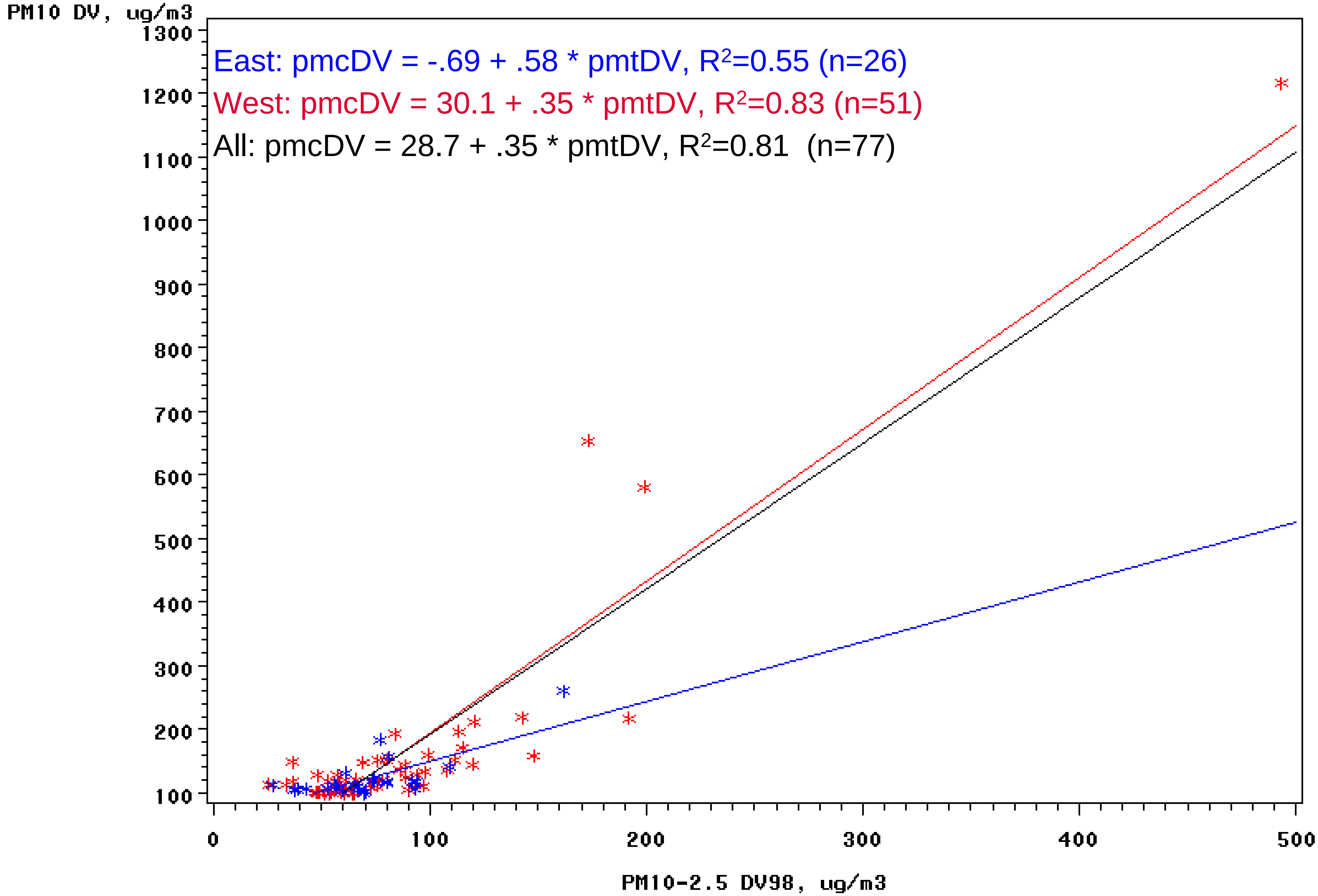


Site ratio distribution statistics.		Times PMt NAAQS	Equiv. PMc NAAQS
p95_98	0.677596	* 150 =	102
p75_98	0.539216	* 150 =	81
mean98	0.437119	* 150 =	66
med_98	0.429527	* 150 =	64
p25_98	0.329518	* 150 =	49
p05_98	0.216104	* 150 =	32

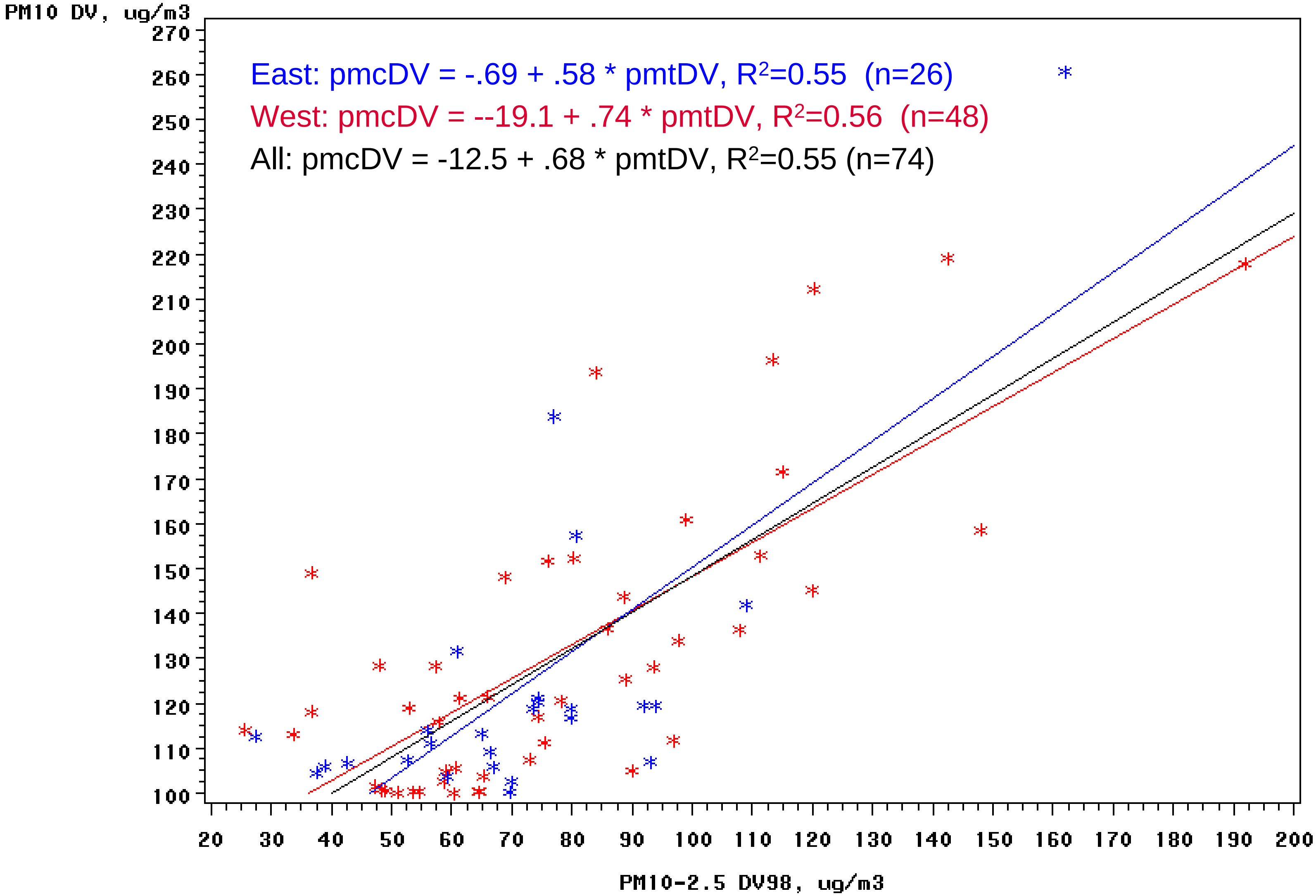
PM10 DV (EE form) versus PM10-2.5 DV (99th percentile form)
all pmt levels

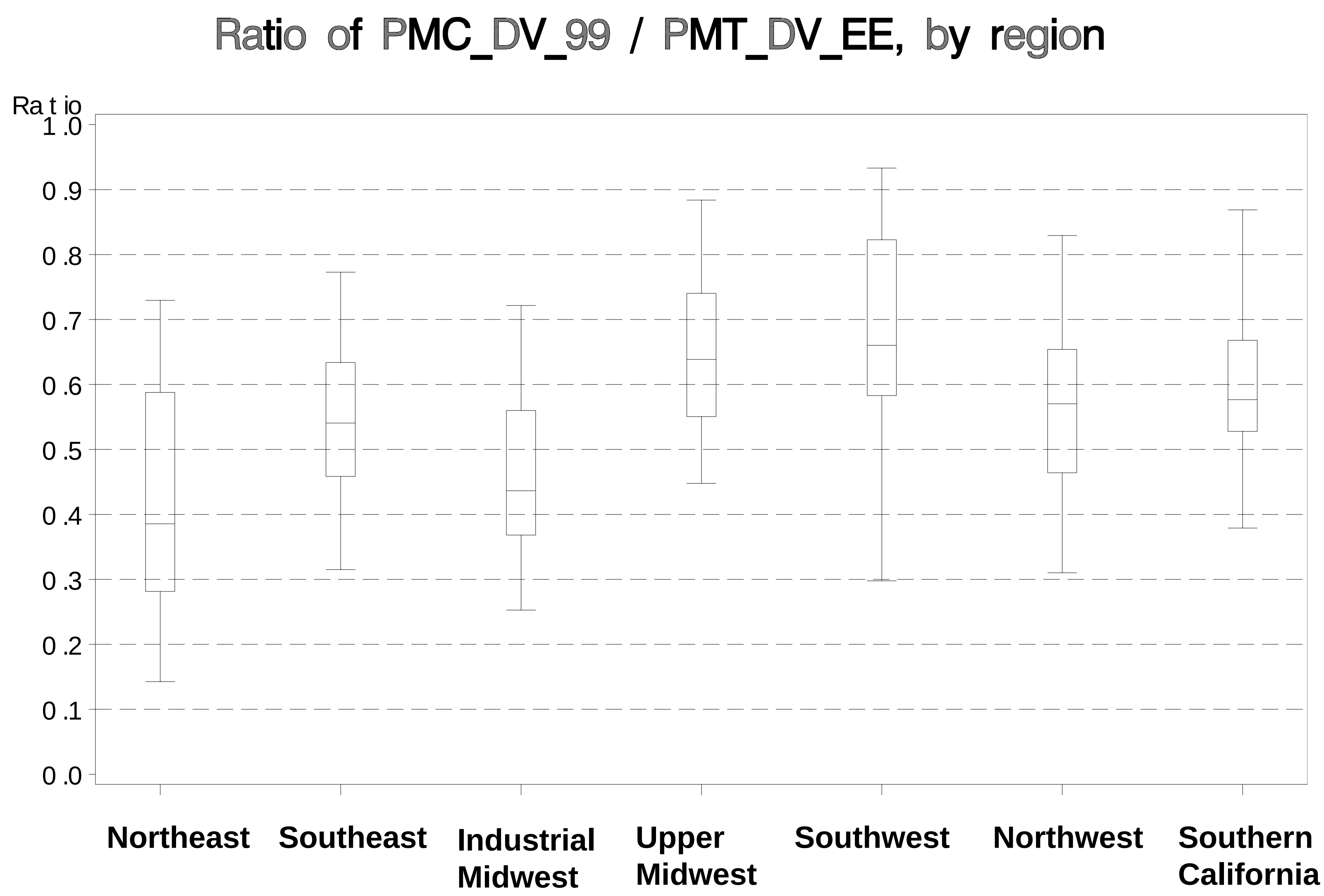


PM10 DV (EE form) versus PM10-2.5 DV (99th percentile form)
pmt ge 100

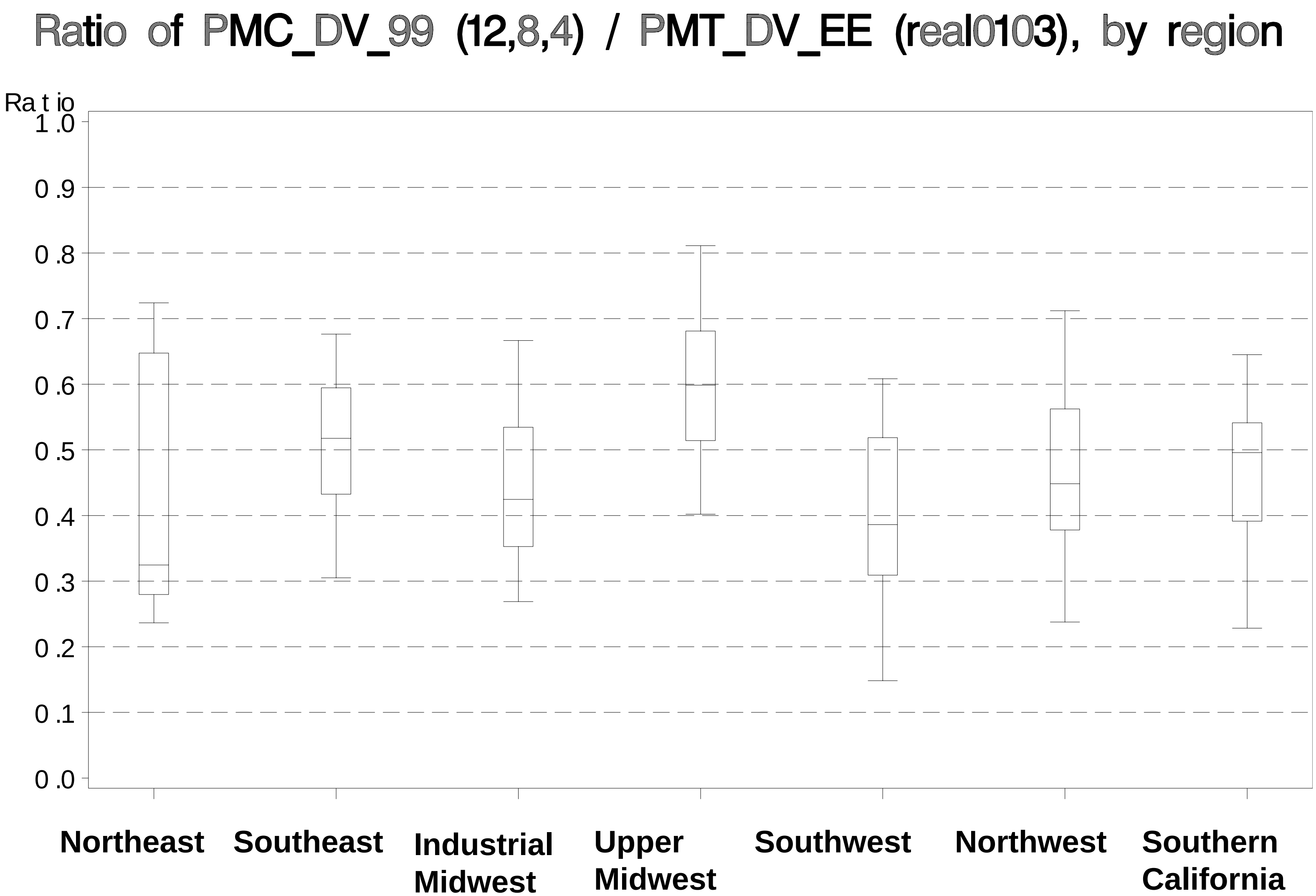


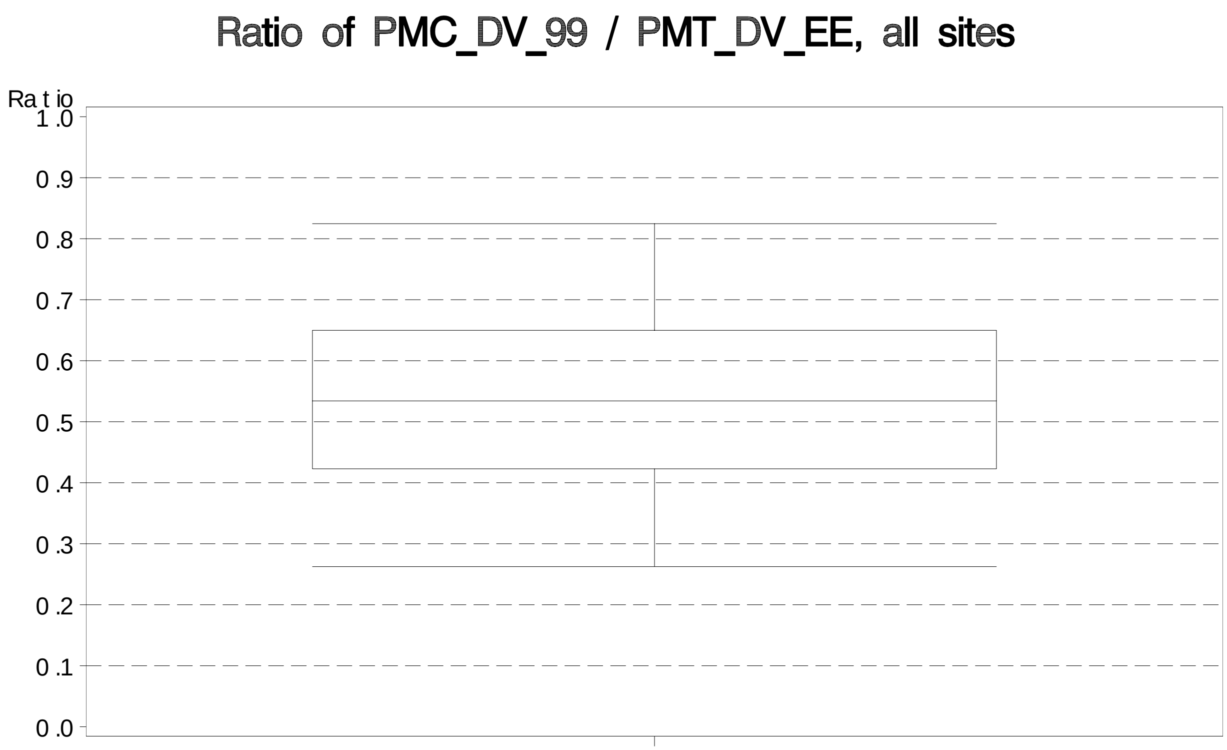
PM10 DV (EE form) versus PM10-2.5 DV (99th percentile form)
pmt ge 100, le 300




Ratio of PMC_DV_99 / PMT_DV_EE, by region
Ratio
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Northeast
Southeast
Industrial Midwest
Upper Midwest
Southwest
Northwest
Southern California

For perspective! – Similar to previous page

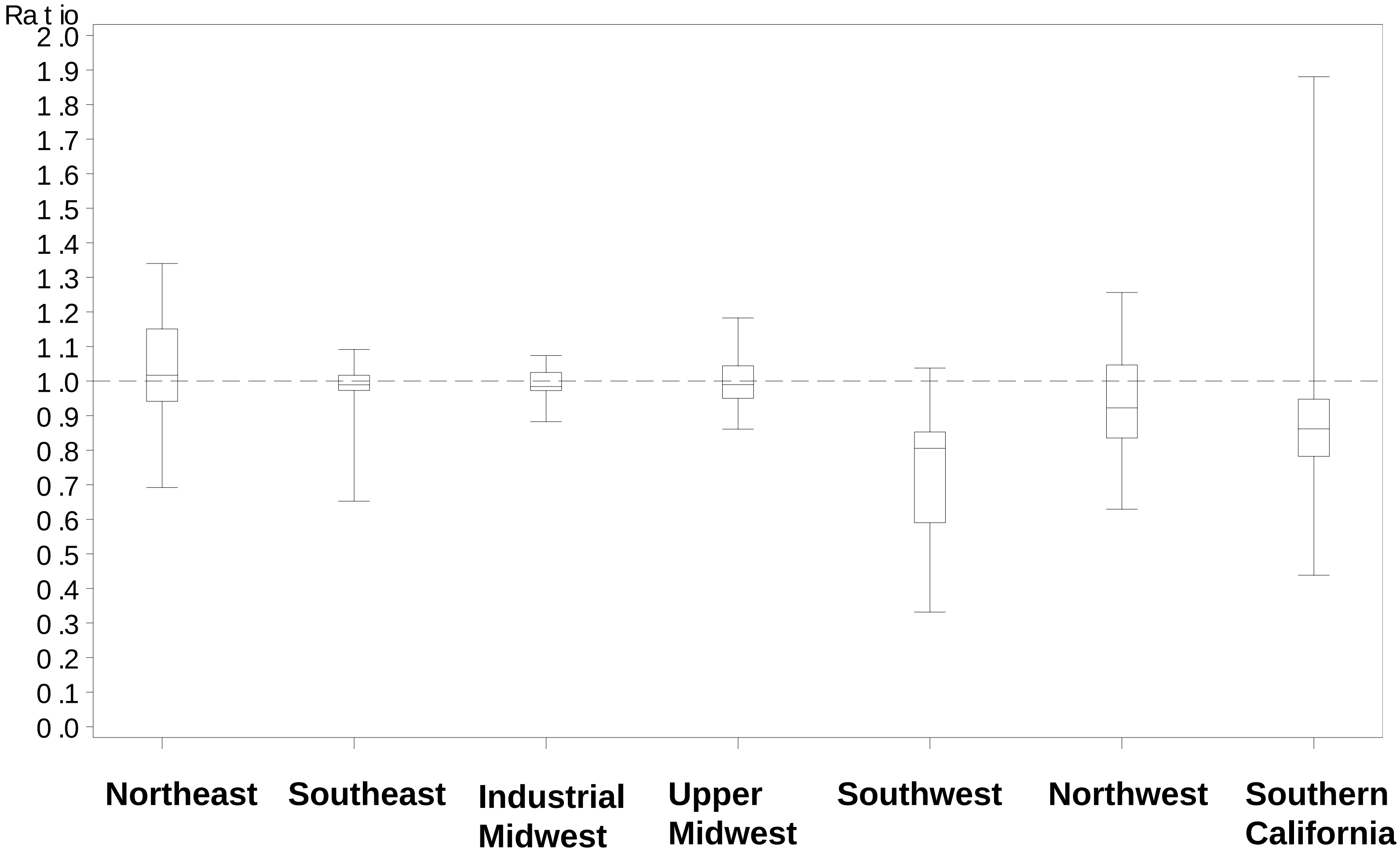




Site ratio distribution statistics.		Times PMt NAAQS	Equiv. PMc NAAQS
p95_99	0.824742	* 150 =	124
p75_99	0.64953	* 150 =	97
mean99	0.538269	* 150 =	81
med_99	0.534156	* 150 =	80
p25_99	0.422535	* 150 =	63
p05_99	0.262361	* 150 =	39

For perspective!

Ratio of PMT_DV_EE (12,8,4) / PMT_DV_EE (12), by region



PM2.5 Spatial Averaging

- Questions:
  - Does spatial averaging allow large populations to go unprotected.
    - I.e., Are there large differences between ‘regular’ (highest site in area) DV’s and spatial average (SA) DV’s?
    - What is the population in areas that could use SA (utilizing current criteria).
    - Would tightening the criteria provide more protection?
  - Are there concerns with spatial averaging?
    - I.e., Are the would-be violating sites in an area that could utilize SA located in lower-income, high percentage-minority, and/or lower education locations than the overall area?
- Analyses details:
  - Started with the default SP PM2.5 database (all sites with 11+ samples in each of the the 12 quarters 2001-2003). Eliminated microscale sites that are not (officially) compared to annual std.
  - Initially enforced the CFR spatial criteria of 1) .6 overall correlation between sites, and 2) no more than 20% difference in site annual mean and spatial annual mean. Did not check criterion that all SA sites should be impacted by similar emissions.
  - Enforced CFR data handling requirement that if SA annual mean is less than or equal the annual std, then only SA sites with 75%+ capture each of the 4 Q’s would have their annual mean included in the spatial annual average (Only 11+ samples required in each of the 4 Q’s if the spatial annual mean was greater than the annual std.) Changed level of std (and completeness check) from 15 to 14 for accurate evaluation of SA effect on those std levels.
  - Constructed SA set of sites by initially considering all sites in the area (CSA, CBSA, or STCOU). If a site-pair correlation was less than cutoff, the lower DV site was eliminated. If e remaining set did not meet annual mean difference criterion then lowest DV site was omitted from set and revised set tried. Continued until reduced set of sites met criteria or less than 2 sites left. Note: Undoubtedly, different combinations of sites (selected w/ rationale and/or at random) could/would meet criteria and yield different results.
  - Only considered (for SA) areas with a regular DV > annual std. level <u>and</u> spatial DV $\geq$ any (valid) non-SA site DV in the area
  - Evaluated appropriateness of .6 (correlation) and 20% (max difference in annual means) levels
  - Tightened the correlation criterion to .9 and the annual mean difference criterion to 10% to evaluate changes in results.

Statistics for Areas that Qualify for Spatial Averaging;

Current Criteria (.6 corr., 20% diff in means), NAAQS Levels of 15, 14, 13

Using criteria of .6 correlation and +/- 20 % difference in annual means. Using annual std level of 15.0				Using criteria of .6 correlation and +/-20 % difference in annual means. Using annual std level of 14.0			
		Could use spatial averaging	Could use SA to meet 15.0 annual std			Could use spatial averaging	Could use SA to meet 14.0 annual std
Number of areas		32	10	Number of areas		45	7
Total population		50,645,671	14,254,268	Total population		63,848,777	8,932,198
Area distribution statistics:				Area distribution statistics:			
Difference in area DV's (ug/m3)	mean	1.06	0.84	Difference in area DV's (ug/m3)	mean	1.01	1.21
	max	2.8	1.5		max	2.8	2.1
	p95	2.7	1.5		p95	2.6	2.1
	p75	1.5	1.2		p75	1.5	2.0
	med	0.9	0.8		med	0.8	1.1
	p25	0.5	0.5		p25	0.4	0.6
	p05	0.2	0.2		p05	0.2	0.2
	min	0.2	0.2		min	0.0	0.2

- Under existing criteria (only considering minimum site correlation and maximum difference in annual means) and considering NAAQS levels of 15 and 14, 32-45 metropolitan areas with a combined population of 51-64 million could qualify for spatial averaging (SA). Note that most of these areas would only lower their area DV and still not attain the standard. But, a lower DV would help these areas attain more quickly, and there are also data capture (less stringent) benefits.
- Assuming these areas could pass (required) additional scrutiny, they would lower their areas DV’s by up to 2.8 ug/m3. (Average reduction in area DV = 1 to 1.1 ug/m3)
- 7-10 of these areas would meet the annual std NAAQS (15 or 14 level) with their spatial average when they couldn’t with their regular site-based DV. Average reduction in DV for these areas is .8-1.2 ug/m3. 9-14 million people live in these areas.

See area listings 1 & 2 next…..

Listing 1: Areas that Qualify for Spatial Averaging; Current Criteria (.6 corr., 20% diff in means), NAAQS Level of 15

Area	Pop.	Number of Sites in Area	Number of Sites in CMZ	Design value without SA	Design value with SA	Difference in DV's	Minimum area site DV	Maximum between-site difference in means	Minimum between-site correlation (annual)	High Site Census Tract Information: Percent minority	High Site: Per capita income	High Site: Median Household Income	High Site: Median Family income	High Site: Average Education Level Attained*	Other Site Census Tract(s) Information (avg.): Percent minority	Other Site: Per capita income	Other Site: Median Household Income	Other Site: Median Family income	Other Site: Average Education Level Attained*	Area (CSA/CBSA) Information: Percent minority	Area: Per capita income	Area: Median Household Income	Area: Median Family Income	Area: Average Education Level Attained*
CBSA_Bakersfield, CA	661,645	5	3	21.8	21.0	0.8	20.3	7.9%	0.98	46%	$11,843	$18,777	$22,669	4.5	32%	$15,947	$33,390	$37,965	5.2	38%	$15,780	$38,858	$42,458	5.0
CBSA_Canton-Massillon, OH	406,934	2	2	17.3	16.6	0.7	15.8	5.6%	0.99	9%	$12,577	$24,205	$30,833	4.4	37%	$14,201	$10,457	$25,000	4.2	9%	$20,154	$36,917	$43,005	5.1
CBSA_Charleston, WV	309,635	2	2	17.1	16.3	0.8	15.5	5.3%	0.97	9%	$16,667	$20,929	$32,167	4.7	20%	$28,021	$27,217	$50,690	5.7	7%	$19,090	$29,508	$35,875	5.0
CBSA_Evansville, IN-KY	342,815	3	3	15.5	15.3	0.2	15.2	3.7%	0.96	11%	$12,773	$29,033	$36,446	4.9	8%	$23,162	$31,037	$46,836	5.1	8%	$20,026	$38,956	$46,128	5.1
CBSA_Hagerstown-Martinsburg, MD-WV	222,771	2	2	16.3	15.1	1.2	14.0	11.0%	0.80	17%	$14,688	$25,423	$35,591	4.3	73%	$21,284	$0	$0	8.9	9%	$19,222	$36,997	$42,510	5.1
CBSA_Huntington-Ashland, WV-KY-OH	288,649	3	3	16.6	15.8	0.8	15.0	8.3%	0.87	12%	$4,312	$6,624	$5,357	7.4	3%	$19,748	$32,969	$38,206	4.9	4%	$16,631	$29,341	$36,169	4.9
CBSA_San Diego-Carlsbad-San Marcos, CA	2,813,833	5	5	15.9	15.0	0.9	14.6	16.4%	0.66	46%	$10,278	$21,021	$23,870	5.1	29%	$16,989	$40,702	$46,701	5.1	33%	$22,928	$51,773	$57,106	5.6
CBSA_Weirton-Steubenville, WV-OH	132,008	4	4	17.8	17.0	0.8	16.2	7.8%	0.86	5%	$15,980	$30,000	$40,181	4.8	6%	$17,242	$33,295	$40,576	4.6	5%	$16,909	$32,335	$39,252	4.9
CBSA_Wheeling, WV-OH	153,172	2	2	15.7	15.4	0.3	15.2	1.9%	0.95	1%	$17,077	$31,836	$39,033	4.8	14%	$8,072	$7,663	$23,214	3.7	4%	$16,749	$29,113	$36,899	5.0
CSA_Birmingham-Hoover-Cullman, AL	1,129,721	8	4	18.0	16.0	2.0	14.7	13.5%	0.77	99%	$12,938	$16,995	$23,333	4.3	16%	$21,918	$45,552	$51,117	5.2	28%	$20,390	$36,593	$43,526	5.1
CSA_Chattanooga-Cleveland-Athens, TN-GA	629,561	3	3	15.6	15.4	0.2	15.2	3.6%	0.87	6%	$14,092	$23,713	$29,183	4.1	24%	$13,257	$22,338	$35,768	4.2	14%	$19,278	$33,613	$39,509	5.0
CSA_Chicago-Naperville-Michigan City, IL-IN-WI	9,312,255	28	2	17.7	17.5	0.2	17.3	4.6%	0.84	10%	$12,368	$31,156	$30,189	4.7	29%	$20,950	$45,553	$53,509	5.1	33%	$24,491	$52,263	$59,135	5.4
CSA_Cincinnati-Middletown-Wilmington, OH-KY-I	2,050,175	12	12	17.8	16.0	1.8	14.5	13.9%	0.90	13%	$19,121	$27,364	$36,667	4.4	22%	$17,950	$31,444	$38,807	4.9	14%	$22,786	$43,248	$49,355	5.3
CSA_Cleveland-Akron-Elyria, OH	2,945,831	13	11	18.3	15.9	2.4	14.2	19.2%	0.84	31%	$15,270	$25,221	$26,850	5.6	41%	$15,278	$28,755	$32,732	4.7	21%	$22,321	$46,452	$53,471	5.3
CSA_Columbus-Auburn-Opelika, GA-AL	420,965	3	3	15.3	14.6	0.7	14.3	8.2%	0.78	65%	$7,295	$10,121	$11,949	3.3	78%	$11,574	$18,636	$23,013	4.2	42%	$17,184	$31,978	$37,256	5.1
CSA_Columbus-Marion-Chillicothe, OH	1,835,189	3	3	16.7	16.2	0.5	15.9	5.9%	0.95	88%	$14,293	$21,486	$27,560	3.9	39%	$15,184	$28,309	$30,408	4.7	17%	$22,256	$45,186	$51,028	5.5
CSA_Dayton-Springfield-Greenville, OH	1,085,094	3	3	15.2	14.7	0.5	14.7	8.5%	0.93	6%	$17,457	$32,708	$40,117	5.3	7%	$16,186	$26,815	$34,558	4.9	16%	$21,263	$42,919	$49,338	5.3
CSA_Detroit-Warren-Flint, MI	5,357,538	14	6	19.5	16.8	2.7	15.1	18.5%	0.83	29%	$7,573	$19,713	$24,031	3.9	43%	$17,486	$35,422	$40,804	5.1	27%	$24,353	$53,256	$60,632	5.4
CSA_Fairmont-Clarksburg, WV	148,742	2	2	15.4	14.7	0.7	14.0	5.6%	0.96	3%	$13,328	$21,839	$28,906	4.6	8%	$14,417	$16,590	$30,031	4.3	4%	$16,094	$28,602	$34,255	4.9
CSA_Fresno-Madera, CA	922,516	2	2	19.7	19.5	0.2	19.2	3.1%	0.97	45%	$12,781	$31,131	$34,440	4.6	57%	$10,976	$16,842	$20,804	4.0	45%	$15,388	$36,870	$39,680	4.7
CSA_Greensboro--Winston-Salem--High Point, N	1,283,856	4	4	15.8	14.6	1.2	14.0	8.6%	0.93	50%	$19,691	$28,094	$34,320	4.6	42%	$25,501	$35,913	$47,006	5.1	25%	$21,090	$38,066	$45,213	5.0
CSA_Harrisburg-Carlisle-Lebanon, PA	629,401	2	2	15.8	14.9	0.9	15.8	13.5%	0.92	35%	$15,752	$31,557	$37,679	4.9	1%	$18,897	$44,341	$50,259	5.1	12%	$21,939	$42,855	$50,094	5.3
CSA_Indianapolis-Anderson-Columbus, IN	1,843,588	6	6	16.7	15.2	1.5	13.6	12.0%	0.92	40%	$9,869	$18,988	$20,417	4.4	17%	$18,785	$36,313	$41,702	5.3	16%	$22,715	$46,925	$53,537	5.4
CSA_Knoxville-Sevierville-La Follette, TN	779,013	5	5	16.7	15.6	1.1	14.2	11.4%	0.85	35%	$7,364	$11,305	$13,239	3.4	11%	$17,905	$35,858	$42,976	5.0	8%	$20,034	$33,904	$40,386	5.1
CSA_Lexington-Fayette--Frankfort--Richmond, K'	602,773	4	4	15.7	14.3	1.4	13.5	9.3%	0.75	28%	$10,418	$17,111	$18,679	5.6	30%	$17,721	$28,083	$36,300	4.8	12%	$20,520	$37,223	$43,417	5.3
CSA_Louisville-Elizabethtown-Scottsburg, KY-IN	1,292,482	6	6	16.9	15.6	1.3	14.1	12.4%	0.85	11%	$13,959	$25,315	$35,469	4.4	11%	$17,611	$27,800	$33,539	4.7	16%	$20,919	$41,171	$46,815	5.2
CSA_Philadelphia-Camden-Vineland, PA-NJ-DE-	5,833,585	14	14	16.4	14.9	1.5	14.3	13.3%	0.90	14%	$42,815	$42,000	$83,904	7.4	31%	$20,897	$40,182	$46,803	5.3	28%	$23,807	$51,473	$59,295	5.3
CSA_Pittsburgh-New Castle, PA	2,525,730	13	3	21.2	18.4	2.8	16.9	17.6%	0.79	2%	$19,491	$35,264	$42,857	4.9	16%	$16,873	$30,404	$38,243	4.8	10%	$20,635	$35,540	$43,510	5.2
CSA_St. Louis-St. Charles-Farmington, MO-IL	2,777,132	12	3	17.5	16.3	1.2	15.2	11.7%	0.79	6%	$17,556	$33,045	$37,313	4.8	37%	$24,136	$39,416	$47,776	5.0	21%	$22,267	$40,513	$47,145	5.3
CSA_Toledo-Fremont, OH	720,980	3	3	15.1	14.9	0.2	14.7	5.3%	0.94	94%	$6,662	$10,171	$10,104	2.9	33%	$14,752	$25,944	$32,969	5.0	16%	$20,529	$41,666	$49,237	5.3
CSA_York-Hanover-Gettysburg, PA	473,043	2	2	17.3	15.4	1.9	13.5	16.1%	0.83	3%	$21,145	$39,962	$47,045	5.3	7%	$18,471	$43,979	$47,042	5.0	7%	$20,603	$43,604	$49,414	5.1
CSA_Youngstown-Warren-East Liverpool, OH-P/	715,039	3	3	15.2	14.8	0.4	14.3	5.1%	0.93	45%	$9,869	$18,150	$30,556	5.3	28%	$16,142	$28,939	$37,757	4.6	12%	$18,399	$34,124	$40,480	5.1

- Areas that could use SA to meet NAAQS are underlined.
- Socioeconomic data from 2000 Census.
- Education Level defined as follows
  - Focused on 'education level attained' (left/lower column)
  - Created 'education average' variable as follows (right/lower formula):
    - (Weighted populations of each category)

```
55. P037001      : pop_mf - Total: Population 25+
  56. P037002      : Male 25+:
  57. P037003      : pop_m1 - Male No schooling completed
  58. P037004      : pop_m2 - Male Nursery-4th grade
  59. P037005      : pop_m3 - Male 5th and 6th grade
  60. P037006      : pop_m4 - Male 7th and 8th grade
  61. P037007      : pop_m5 - Male 9th grade
  62. P037008      : pop_m6 - Male 10th grade
  63. P037009      : pop_m7 - Male 11th grade
  64. P037010      : pop_m8 - Male 12th grade, no diploma
  65. P037011      : pop_m9 - Male High school grad (inc equivalency)
  66. P037012      : pop_m10 - Male Some college, under 1 year
  67. P037013      : pop_m11 - Male Some college, 1+ years, no degree
  68. P037014      : pop_m12 - Male Associate degree
  69. P037015      : pop_m13 - Male Bachelor's degree
  70. P037016      : pop_m14 - Male Master's degree
  71. P037017      : pop_m15 - Male Professional school degree
  72. P037018      : pop_m16 -Male Doctorate degree
  73. P037019      : Female 25+:
  74. P037020      : pop_f1 - Female No schooling completed
  75. P037021      : pop_f2 - Female Nursery-4th grade
  76. P037022      : pop_f3 - Female 5th and 6th grade
  77. P037023      : pop_f4 - Female 7th and 8th grade
  78. P037024      : pop_f5 - Female 9th grade
  79. P037025      : pop_f6 - Female 10th grade
  80. P037026      : pop_f7 - Female 11th grade
  81. P037027      : pop_f8 - Female 12th grade, no diploma
  82. P037028      : pop_f9 - Female High school grad (inc equivalency)
  83. P037029      : pop_f10 - Female Some college, under 1 year
  84. P037030      : pop_f11 - Female Some college, 1+ years, no degree
  85. P037031      : pop_f12 - Female Associate degree
  86. P037032      : pop_f13 - Female Bachelor's degree
  87. P037033      : pop_f14 - Female Master's degree
  88. P037034      : pop_f15 - Female Professional school degree
  89. P037035      : pop_f16 - Female Doctorate degree
```

```
avg_ed=
((pop_m1+pop_f1*1)+(pop_m2+pop_f2*2)+
(pop_m3+pop_f3*3)+(pop_m4+pop_f4*4)+
(pop_m5+pop_f5*5)+(pop_m6+pop_f6*6)+
(pop_m7+pop_f7*7)+(pop_m8+pop_f8*8)+
(pop_m9+pop_f9*9)+(pop_m10+pop_f10*10)+
(pop_m11+pop_f11*11)+(pop_m12+pop_f12*12)+
(pop_m13+pop_f13*13)+(pop_m14+pop_f14*14)+
(pop_m15+pop_f15*15)+(pop_m16+pop_f16*16))
/pop_mf;
```

Listing 2: Areas that Qualify for Spatial Averaging; Current Criteria (.6 corr., 20% diff in means), NAAQS Level of 14

										High Site Census Tract Information					Other Site Census Tract(s) Information (avg.)					Area (CSA/CBSA) Information				
Area	Pop.	Number of Sites in Area	Number of Sites in CMZ	Design value without SA	Design value with SA	Difference in DV's	Minimum area site DV	Maximum between-site difference in means	Minimum between-site correlation (annual)	Percent minority	Per capita income	Median Household Income	Median Family income	Average Education Level Attained*	Percent minority	Per capita income	Median Household Income	Median Family income	Average Education Level Attained*	Percent minority	Per capita income	Median Household Income	Median Family Income	Average Education Level Attained*
CBSA_Allentown-Bethlehem-Easton, PA-NJ	740,395	3	3	14.8	14.4	0.4	14.6	4.8%	0.91	12%	$17,983	$44,297	$48,333	4.8	7%	$21,244	$34,187	$48,800	4.6	10%	$21,867	$44,922	$52,674	5.2
CBSA_Augusta-Richmond County, GA-SC	499,684	2	2	14.7	13.2	1.5	12.4	9.6%	0.84	40%	$14,902	$29,783	$32,813	4.4	53%	$17,757	$36,991	$40,950	4.9	39%	$18,496	$37,529	$43,751	5.1
CBSA_Bakersfield, CA	661,645	5	3	21.8	21.0	0.8	20.3	7.9%	0.98	46%	$11,843	$18,777	$22,669	4.5	32%	$15,947	$33,390	$37,965	5.2	38%	$15,780	$38,858	$42,458	5.0
CBSA_Canton-Massillon, OH	406,934	2	2	17.3	16.6	0.7	15.8	5.6%	0.99	9%	$12,577	$24,205	$30,833	4.4	37%	$14,201	$10,457	$25,000	4.2	9%	$20,154	$36,917	$43,005	5.1
CBSA_Charleston, WV	309,635	2	2	17.1	16.3	0.8	15.5	5.3%	0.97	9%	$16,667	$20,929	$32,167	4.7	20%	$28,021	$27,217	$50,690	5.7	7%	$19,090	$29,508	$35,875	5.0
CBSA_Evansville, IN-KY	342,815	3	3	15.5	15.3	0.2	15.2	3.7%	0.96	11%	$12,773	$29,033	$36,446	4.9	8%	$23,162	$31,037	$46,836	5.1	8%	$20,026	$38,956	$46,128	5.1
CBSA_Hagerstown-Martinsburg, MD-WV	222,771	2	2	16.3	15.1	1.2	14.0	11.0%	0.80	17%	$14,688	$25,423	$35,591	4.3	73%	$21,284	$0	$0	8.9	9%	$19,222	$36,997	$42,510	5.1
CBSA_Huntington-Ashland, WV-KY-OH	288,649	3	3	16.6	15.8	0.8	15.0	8.3%	0.87	12%	$4,312	$6,624	$5,357	7.4	3%	$19,748	$32,969	$38,206	4.9	4%	$16,631	$29,341	$36,169	4.9
CBSA_Roanoke, VA	288,309	2	2	14.7	14.4	0.3	14.2	2.4%	0.96	14%	$15,721	$29,774	$37,699	4.7	6%	$22,330	$41,331	$50,891	5.5	15%	$21,006	$38,681	$45,437	5.1
CBSA_San Diego-Carlsbad-San Marcos, CA	2,813,833	5	5	15.9	15.0	0.9	14.6	16.4%	0.66	46%	$10,278	$21,021	$23,870	5.1	29%	$16,989	$40,702	$46,701	5.1	33%	$22,928	$51,773	$57,106	5.6
CBSA_South Bend-Mishawaka, IN-MI	316,663	3	3	14.3	14.1	0.2	14.0	2.6%	0.99	64%	$12,615	$25,466	$27,993	4.5	36%	$14,681	$32,046	$35,594	4.8	17%	$19,728	$39,967	$45,577	5.3
CBSA_Terre Haute, IN	170,943	2	2	14.6	14.0	0.6	13.4	6.2%	0.96	7%	$16,572	$32,321	$39,474	5.2	5%	$19,748	$38,281	$45,710	5.1	7%	$17,342	$35,029	$41,115	5.2
CBSA_Weirton-Steubenville, WV-OH	132,008	4	4	17.8	17.0	0.8	16.2	7.8%	0.86	5%	$15,980	$30,000	$40,181	4.8	6%	$17,242	$33,295	$40,576	4.6	5%	$16,909	$32,335	$39,252	4.9
CBSA_Wheeling, WV-OH	153,172	2	2	15.7	15.4	0.3	15.2	1.9%	0.95	1%	$17,077	$31,836	$39,033	4.8	14%	$8,072	$7,663	$23,214	3.7	4%	$16,749	$29,113	$36,899	5.0
CSA_Birmingham-Hoover-Cullman, AL	1,129,721	8	6	18.0	15.4	2.6	13.8	18.1%	0.77	99%	$12,938	$16,995	$23,333	4.3	16%	$21,918	$45,552	$51,117	5.2	28%	$20,390	$36,593	$43,526	5.1
CSA_Charlotte-Gastonia-Salisbury, NC-SC	1,897,034	5	5	14.9	14.4	0.5	14.0	6.1%	0.92	92%	$12,094	$26,829	$28,413	4.5	27%	$20,137	$37,554	$44,614	5.0	26%	$22,291	$39,740	$45,842	5.2
CSA_Chattanooga-Cleveland-Athens, TN-GA	629,561	3	3	15.6	15.4	0.2	14.6	3.6%	0.87	6%	$14,092	$23,713	$29,183	4.1	24%	$13,257	$22,338	$35,768	4.2	14%	$19,278	$33,613	$39,509	5.0
CSA_Chicago-Naperville-Michigan City, IL-IN-WI	9,312,255	28	2	17.7	17.5	0.2	17.3	4.6%	0.84	10%	$12,368	$31,156	$30,189	4.7	29%	$20,950	$45,553	$53,509	5.1	33%	$24,491	$52,263	$59,135	5.4
CSA_Cincinnati-Middletown-Wilmington, OH-KY-I	2,050,175	12	12	17.8	16.0	1.8	14.5	13.9%	0.90	13%	$19,121	$27,364	$36,667	4.4	22%	$17,950	$31,444	$38,807	4.9	14%	$22,786	$43,248	$49,355	5.3
CSA_Cleveland-Akron-Elyria, OH	2,945,831	13	11	18.3	15.9	2.4	14.2	19.2%	0.84	31%	$15,270	$25,221	$26,850	5.6	41%	$15,278	$28,755	$32,732	4.7	21%	$22,321	$46,452	$53,471	5.3
CSA_Columbus-Auburn-Opelika, GA-AL	420,965	3	3	15.3	14.6	0.7	14.3	8.2%	0.78	65%	$7,295	$10,121	$11,949	3.3	78%	$11,574	$18,636	$23,013	4.2	42%	$17,184	$31,978	$37,256	5.1
CSA_Columbus-Marion-Chillicothe, OH	1,835,189	3	3	16.7	16.2	0.5	15.9	5.9%	0.95	88%	$14,293	$21,486	$27,560	3.9	39%	$15,184	$28,309	$30,408	4.7	17%	$22,256	$45,186	$51,028	5.5
CSA_Dayton-Springfield-Greenville, OH	1,085,094	3	3	15.2	14.7	0.5	14.7	8.5%	0.93	6%	$17,457	$32,708	$40,117	5.3	7%	$16,186	$26,815	$34,558	4.9	16%	$21,263	$42,919	$49,338	5.3
CSA_Detroit-Warren-Flint, MI	5,357,538	14	6	19.5	16.8	2.7	15.1	18.5%	0.83	29%	$7,573	$19,713	$24,031	3.9	43%	$17,486	$35,422	$40,804	5.1	27%	$24,353	$53,256	$60,632	5.4
CSA_Fairmont-Clarksburg, WV	148,742	2	2	15.4	14.7	0.7	14.0	5.6%	0.96	3%	$13,328	$21,839	$28,906	4.6	8%	$14,417	$16,590	$30,031	4.3	4%	$16,094	$28,602	$34,255	4.9
CSA_Fort Wayne-Huntington-Auburn, IN	548,416	2	2	14.3	14.3	0.0	14.3	1.1%	0.99	14%	$15,132	$24,423	$36,659	4.3	13%	$19,343	$39,929	$44,730	5.4	11%	$20,468	$43,571	$49,877	5.3
CSA_Fresno-Madera, CA	922,516	2	2	19.7	19.5	0.2	19.2	3.1%	0.97	45%	$12,781	$31,131	$34,440	4.6	57%	$10,976	$16,842	$20,804	4.0	45%	$15,388	$36,870	$39,680	4.7
CSA_Greensboro--Winston-Salem--High Point, N	1,283,856	4	4	15.8	14.6	1.2	14.0	8.6%	0.93	50%	$19,691	$28,094	$34,320	4.6	42%	$25,501	$35,913	$47,006	5.1	25%	$21,090	$38,066	$45,213	5.0
CSA_Greenville-Anderson-Seneca, SC	791,895	2	2	14.5	12.5	2.0	10.6	18.1%	0.88	11%	$20,873	$47,161	$54,688	5.4	2%	$16,573	$30,429	$36,127	4.9	19%	$19,843	$36,301	$43,552	5.1
CSA_Harrisburg-Carlisle-Lebanon, PA	629,401	2	2	15.8	14.9	0.9	15.8	13.5%	0.92	35%	$15,752	$31,557	$37,679	4.9	1%	$18,897	$44,341	$50,259	5.1	12%	$21,939	$42,855	$50,094	5.3
CSA_Houston-Baytown-Huntsville, TX	4,815,122	6	6	14.2	12.1	2.1	10.9	19.4%	0.61	97%	$10,236	$24,353	$24,457	4.0	45%	$15,390	$38,444	$42,128	4.7	37%	$21,519	$41,701	$47,600	5.3
CSA_Huntsville-Decatur, AL	488,243	2	2	14.1	13.9	0.2	13.7	2.7%	0.88	54%	$13,252	$17,589	$23,000	5.2	12%	$19,520	$48,507	$54,079	4.2	23%	$21,033	$38,629	$45,429	5.4
CSA_Indianapolis-Anderson-Columbus, IN	1,843,588	6	6	16.7	15.2	1.5	13.6	12.0%	0.92	40%	$9,869	$18,988	$20,417	4.4	17%	$18,785	$36,313	$41,702	5.3	16%	$22,715	$46,925	$53,537	5.4
CSA_Johnson City-Kingsport-Bristol, TN-VA	480,091	2	2	14.7	14.5	0.2	14.3	2.4%	0.96	4%	$18,538	$25,522	$31,715	4.3	12%	$15,781	$24,412	$27,723	4.0	4%	$17,800	$31,032	$37,582	4.9
CSA_Knoxville-Sevierville-La Follette, TN	779,013	5	5	16.7	15.6	1.1	14.2	11.4%	0.85	35%	$7,364	$11,305	$13,239	3.4	11%	$17,905	$35,858	$42,976	5.0	8%	$20,034	$33,904	$40,386	5.1
CSA_Lexington-Fayette--Frankfort--Richmond, K'	602,773	4	4	15.7	14.3	1.4	13.5	9.3%	0.75	28%	$10,418	$17,111	$18,679	5.6	30%	$17,721	$28,083	$36,300	4.8	12%	$20,520	$37,223	$43,417	5.3
CSA_Little Rock-North Little Rock-Pine Bluff, AR	785,024	5	5	14.1	13.0	1.1	11.9	13.0%	0.78	89%	$8,205	$18,099	$21,758	4.0	16%	$15,474	$33,680	$40,409	5.3	26%	$19,069	$35,771	$41,537	5.2
CSA_Louisville-Elizabethtown-Scottsburg, KY-IN	1,292,482	6	6	16.9	15.6	1.3	14.1	12.4%	0.85	11%	$13,959	$25,315	$35,469	4.4	11%	$17,611	$27,800	$33,539	4.7	16%	$20,919	$41,171	$46,815	5.2
CSA_Nashville-Davidson--Murfreesboro--Columb	1,381,287	3	3	14.4	13.4	1.0	13.5	7.4%	0.88	23%	$20,803	$40,781	$49,598	5.3	11%	$21,017	$41,519	$50,386	5.1	20%	$22,287	$42,067	$48,075	5.3
CSA_Philadelphia-Camden-Vineland, PA-NJ-DE-	5,833,585	14	14	16.4	14.9	1.5	14.3	13.3%	0.90	14%	$42,815	$42,000	$83,904	7.4	31%	$20,897	$40,182	$46,803	5.3	28%	$23,807	$51,473	$59,295	5.3
CSA_Pittsburgh-New Castle, PA	2,525,730	13	3	21.2	18.4	2.8	16.9	17.6%	0.79	2%	$19,491	$35,264	$42,857	4.9	16%	$16,873	$30,404	$38,243	4.8	10%	$20,635	$35,540	$43,510	5.2
CSA_St. Louis-St. Charles-Farmington, MO-IL	2,777,132	12	9	17.5	15.3	2.2	14.5	19.2%	0.76	6%	$17,556	$33,045	$37,313	4.8	37%	$24,136	$39,416	$47,776	5.0	21%	$22,267	$40,513	$47,145	5.3
CSA_Toledo-Fremont, OH	720,980	3	3	15.1	14.9	0.2	14.7	5.3%	0.94	94%	$6,662	$10,171	$10,104	2.9	33%	$14,752	$25,944	$32,969	5.0	16%	$20,529	$41,666	$49,237	5.3
CSA_York-Hanover-Gettysburg, PA	473,043	2	2	17.3	15.4	1.9	13.5	16.1%	0.83	3%	$21,145	$39,962	$47,045	5.3	7%	$18,471	$43,979	$47,042	5.0	7%	$20,603	$43,604	$49,414	5.1
CSA_Youngstown-Warren-East Liverpool, OH-P/	715,039	3	3	15.2	14.8	0.4	14.3	5.1%	0.93	45%	$9,869	$18,150	$30,556	5.3	28%	$16,142	$28,939	$37,757	4.6	12%	$18,399	$34,124	$40,480	5.1

- Areas that could use SA to meet NAAQS are underlined.

Concerns w/ Spatial Averaging

- Are the would-be violating ('high') sites in an area that could use SA located in lower-income, high percentage-minority, and/or lower education locations than the overall area?

Comparison of High-Site Census Tract Socioeconomic Data to Area Average

NAAQS Level of 15

	Areas that could use spatial averaging			Areas that could attain the standard using spatial averaging (subset of left columns)		
Variable	Total	Number where indicated metric is higher for the metro area than in the high-site census tract	Number where indicated metric is lower for the metro area than in the high site census tract	Total	Number where indicated metric is higher for the metro area than in the high-site census tract	Number where indicated metric is lower for the metro area than in the high site census tract
Percentage Minority	32	13	19	10	3	7
Per Capita Income	32	29	3	10	9	1
Median Family Income	32	31	1	10	10	0
Median Household Income	32	29	3	10	9	1
Education Level Attained	32	25	7	10	6	4

NAAQS Level of 14

	Areas that could use spatial averaging			averaging (subset of left columns)		
Variable	Total	Number where indicated metric is higher for the metro area than in the high-site census tract	Number where indicated metric is lower for the metro area than in the high site census tract	Total	Number where indicated metric is higher for the metro area than in the high-site census tract	Number where indicated metric is lower for the metro area than in the high site census tract
Percentage Minority	45	15	30	7	1	6
Per Capita Income	45	40	5	7	6	1
Median Family Income	45	43	2	7	6	1
Median Household Income	45	40	5	7	5	2
Education Level Attained	45	36	9	7	5	2

In most areas that could use SA (15 or 14 NAAQS level), the high site is located in an area populated by lower income, higher percentage minority, and less-educated people when compared to the overall metro area.

Concerns w/ Spatial Averaging

- Is there a relationship between the magnitude of the DV disparity and the disparity in the socioeconomic variables?
- See computations below for NAAQS level of 14.

Area	Difference in DV's	Difference Between High-Site Census Tract and Other-Site Census Tract(s)					Difference Between High-Site Census Tract and Area (CSA/CBSA) Average				
		Percent minority	Per capita income	Median Household Income	Median Family Income	Average Education Level Attained*	Percent minority	Per capita income	Median Household Income	Median Family Income	Average Education Level Attained*
CBSA_Allentown-Bethlehem-Easton, PA-NJ	0.4	6%	-$3,261	$10,110	-$467	-0.4	3%	-$3,884	-$625	-$4,341	-0.4
CBSA_Augusta-Richmond County, GA-SC	1.5	-13%	-$2,855	-$7,208	-$8,137	-0.7	1%	-$3,594	-$7,746	-$10,938	-0.7
CBSA_Bakersfield, CA	0.8	14%	-$4,104	-$14,613	-$15,296	-0.5	8%	-$3,937	-$20,081	-$19,789	-0.5
CBSA_Canton-Massillon, OH	0.7	-28%	-$1,624	$13,748	$5,833	-0.7	0%	-$7,577	-$12,712	-$12,172	-0.7
CBSA_Charleston, WV	0.8	-11%	-$11,354	-$6,288	-$18,523	-0.3	2%	-$2,423	-$8,579	-$3,708	-0.3
CBSA_Evansville, IN-KY	0.2	3%	-$10,389	-$2,004	-$10,390	-0.2	3%	-$7,253	-$9,923	-$9,682	-0.2
CBSA_Hagerstown-Martinsburg, MD-WV	1.2	-56%	-$6,596			-0.8	9%	-$4,534	-$11,574	-$6,919	-0.8
CBSA_Huntington-Ashland, WV-KY-OH	0.8	9%	-$15,436	-$26,345	-$32,849	2.5	8%	-$12,319	-$22,717	-$30,812	2.5
CBSA_Roanoke, VA	0.3	9%	-$6,609	-$11,557	-$13,192	-0.4	-1%	-$5,285	-$8,907	-$7,738	-0.4
CBSA_San Diego-Carlsbad-San Marcos, CA	0.9	18%	-$6,711	-$19,681	-$22,832	-0.5	13%	-$12,650	-$30,752	-$33,237	-0.5
CBSA_South Bend-Mishawaka, IN-MI	0.2	28%	-$2,066	-$6,580	-$7,601	-0.8	48%	-$7,113	-$14,501	-$17,584	-0.8
CBSA_Terre Haute, IN	0.6	2%	-$3,176	-$5,960	-$6,236	0.0	0%	-$770	-$2,708	-$1,641	0.0
CBSA_Weirton-Steubenville, WV-OH	0.8	-1%	-$1,262	-$3,295	-$395	-0.2	-1%	-$929	-$2,335	$929	-0.2
CBSA_Wheeling, WV-OH	0.3	-13%	$9,005	$24,173	$15,819	-0.2	-3%	$328	$2,723	$2,134	-0.2
CSA_Birmingham-Hoover-Cullman, AL	2.6	83%	-$8,980	-$28,557	-$27,784	-0.8	70%	-$7,452	-$19,598	-$20,193	-0.8
CSA_Charlotte-Gastonia-Salisbury, NC-SC	0.5	65%	-$8,043	-$10,725	-$16,201	-0.7	66%	-$10,197	-$12,911	-$17,429	-0.7
CSA_Chattanooga-Cleveland-Athens, TN-GA	0.2	-19%	$835	$1,376	-$6,585	-0.8	-9%	-$5,186	-$9,900	-$10,326	-0.8
CSA_Chicago-Naperville-Michigan City, IL-IN-WI	0.2	-19%	-$8,582	-$14,397	-$23,320	-0.7	-23%	-$12,123	-$21,107	-$28,946	-0.7
CSA_Cincinnati-Middletown-Wilmington, OH-KY-I	1.8	-9%	$1,171	-$4,080	-$2,140	-0.9	-1%	-$3,665	-$15,884	-$12,688	-0.9
CSA_Cleveland-Akron-Elyria, OH	2.4	-10%	-$8	-$3,534	-$5,882	0.3	10%	-$7,051	-$21,231	-$26,621	0.3
CSA_Columbus-Auburn-Opelika, GA-AL	0.7	-13%	-$4,279	-$8,515	-$11,064	-1.8	23%	-$9,889	-$21,857	-$25,307	-1.8
CSA_Columbus-Marion-Chillicothe, OH	0.5	49%	-$891	-$6,823	-$2,848	-1.6	72%	-$7,963	-$23,700	-$23,468	-1.6
CSA_Dayton-Springfield-Greenville, OH	0.5	-1%	$1,271	$5,894	$5,560	0.1	-10%	-$3,806	-$10,211	-$9,221	0.1
CSA_Detroit-Warren-Flint, MI	2.7	-14%	-$9,913	-$15,709	-$16,773	-1.5	1%	-$16,780	-$33,543	-$36,601	-1.5
CSA_Fairmont-Clarksburg, WV	0.7	-5%	-$1,089	$5,249	-$1,125	-0.3	-1%	-$2,766	-$6,763	-$5,349	-0.3
CSA_Fort Wayne-Huntington-Auburn, IN	0.0	1%	-$4,211	-$15,506	-$8,071	-1.0	3%	-$5,336	-$19,148	-$13,218	-1.0
CSA_Fresno-Madera, CA	0.2	-12%	$1,805	$14,289	$13,636	-0.1	1%	-$2,607	-$5,739	-$5,240	-0.1
CSA_Greensboro--Winston-Salem--High Point, N	1.2	7%	-$5,810	-$7,819	-$12,686	-0.4	25%	-$1,399	-$9,972	-$10,893	-0.4
CSA_Greenville-Anderson-Seneca, SC	2.0	10%	$4,300	$16,732	$18,561	0.3	-8%	$1,030	$10,860	$11,136	0.3
CSA_Harrisburg-Carlisle-Lebanon, PA	0.9	34%	-$3,145	-$12,784	-$12,580	-0.3	23%	-$6,187	-$11,298	-$12,415	-0.3
CSA_Houston-Baytown-Huntsville, TX	2.1	52%	-$5,154	-$14,091	-$17,671	-1.3	60%	-$11,283	-$17,348	-$23,143	-1.3
CSA_Huntsville-Decatur, AL	0.2	42%	-$6,268	-$30,918	-$31,079	-0.2	31%	-$7,781	-$21,040	-$22,429	-0.2
CSA_Indianapolis-Anderson-Columbus, IN	1.5	23%	-$8,916	-$17,325	-$21,285	-1.0	24%	-$12,846	-$27,937	-$33,120	-1.0
CSA_Johnson City-Kingsport-Bristol, TN-VA	0.2	-8%	$2,757	$1,110	$3,992	-0.5	0%	$738	-$5,510	-$5,867	-0.5
CSA_Knoxville-Sevierville-La Follette, TN	1.1	24%	-$10,541	-$24,553	-$29,737	-1.8	27%	-$12,670	-$22,599	-$27,147	-1.8
CSA_Lexington-Fayette--Frankfort--Richmond, K	1.4	-2%	-$7,303	-$10,972	-$17,621	0.3	16%	-$10,102	-$20,112	-$24,738	0.3
CSA_Little Rock-North Little Rock-Pine Bluff, AR	1.1	74%	-$7,269	-$15,581	-$18,651	-1.2	63%	-$10,864	-$17,672	-$19,779	-1.2
CSA_Louisville-Elizabethtown-Scottsburg, KY-IN	1.3	0%	-$3,652	-$2,485	$1,930	-0.8	-5%	-$6,960	-$15,856	-$11,346	-0.8
CSA_Nashville-Davidson--Murfreesboro--Columb	1.0	12%	-$214	-$738	-$788	0.0	3%	-$1,484	-$1,286	$1,523	0.0
CSA_Philadelphia-Camden-Vineland, PA-NJ-DE-	1.5	-17%	$21,918	$1,818	$37,101	2.1	-14%	$19,008	-$9,473	$24,609	2.1
CSA_Pittsburgh-New Castle, PA	2.8	-14%	$2,618	$4,860	$4,614	-0.3	-8%	-$1,144	-$276	-$653	-0.3
CSA_St. Louis-St. Charles-Farmington, MO-IL	2.2	-32%	-$6,580	-$6,371	-$10,463	-0.5	-15%	-$4,711	-$7,468	-$9,832	-0.5
CSA_Toledo-Fremont, OH	0.2	61%	-$8,090	-$15,773	-$22,865	-2.3	78%	-$13,867	-$31,495	-$39,133	-2.3
CSA_York-Hanover-Gettysburg, PA	1.9	-4%	$2,674	-$4,017	$3	0.2	-4%	$542	-$3,642	-$2,369	0.2
CSA_Youngstown-Warren-East Liverpool, OH-PA	0.4	18%	-$6,273	-$10,789	-$7,201	0.2	34%	-$8,530	-$15,974	-$9,924	0.2
Correlation between DV difference column and socioeconomic variable difference columns		**-0.045839**	**0.061794**	**-0.0862136**	**0.022738**	**0.0460458**	**-0.037935**	**0.031241**	**-0.0253302**	**-0.011631**	**0.0460458**

•There does not appear to be a relationship between magnitude of DV disparity and the disparity in the socioeconomic variables.

•There are obviously many other factors that determine differences in the socioeconomic variables <u>across areas.</u>

Concerns w/ Spatial Averaging

- Within an area, is there a relationship between DV level and the socioeconomic variable level?
- Assume other factors cause differences across areas. Look for relationships within areas. Look in all areas with multiple sites, not just areas where SA is applicable.

Correlation of Within-Area Monitoring Site Tract Data - DV versus Percent Minority

Areas with multiple Sites			Areas with 2 Sites			Areas with 3+ Sites				
Number	Number w/ positive correlation	Percent areas w/ negative correlation	Number	Number w/ positive correlation	Percent areas w/ positive correlation	Number	Number w/ positive correlation	Percent areas w/ positive correlation	Mean Correlation (where positive)	Median Correlation (where positive)
125	84	67%	50	24	48%	75	60	80%	0.6175	0.659

Correlation of Within-Area Monitoring Site Tract Data - DV versus Per Capita Income

Areas with multiple Sites			Areas with 2 Sites			Areas with 3+ Sites				
Number	Number w/ negative correlation	Percent areas w/ negative correlation	Number	Number w/ negative correlation	Percent areas w/ negative correlation	Number	Number w/ negative correlation	Percent areas w/ negative correlation	Mean Correlation (where negative)	Median Correlation (where negative)
125	93	74%	50	35	70%	75	58	77%	-0.5906	-0.5967

Correlation of Within-Area Monitoring Site Tract Data - DV versus Median Household Income

Areas with multiple Sites			Areas with 2 Sites			Areas with 3+ Sites				
Number	Number w/ negative correlation	Percent areas w/ negative correlation	Number	Number w/ negative correlation	Percent areas w/ negative correlation	Number	Number w/ negative correlation	Percent areas w/ negative correlation	Mean Correlation (where negative)	Median Correlation (where negative)
125	92	74%	50	32	64%	75	60	80%	-0.5791	-0.5615

Correlation of Within-Area Monitoring Site Tract Data - DV versus Median Family Income

Areas with multiple Sites			Areas with 2 Sites			Areas with 3+ Sites				
Number	Number w/ negative correlation	Percent areas w/ negative correlation	Number	Number w/ negative correlation	Percent areas w/ negative correlation	Number	Number w/ negative correlation	Percent areas w/ negative correlation	Mean Correlation (where negative)	Median Correlation (where negative)
125	96	77%	50	33	66%	75	63	84%	-0.599	-0.6246

Correlation of Within-Area Monitoring Site Tract Data - DV versus Average Education

Areas with multiple Sites			Areas with 2 Sites			Areas with 3+ Sites				
Number	Number w/ negative correlation	Percent areas w/ negative correlation	Number	Number w/ negative correlation	Percent areas w/ negative correlation	Number	Number w/ negative correlation	Percent areas w/ negative correlation	Mean Correlation (where negative)	Median Correlation (where negative)
125	71	57%	50	27	54%	75	44	59%	-0.5662	-0.5886

•In most areas, there appears to be a negative relationship between DV and 1) education level attained, 2) per capita income, 3) median household income, and 4)median family income

•In most areas, there is a positive relationship between DV and percentage minority.

Is the Tightening of SA Criteria Appropriate?

- The 2 considered SA criterions --- .6 minimum correlation and 20% +/- maximum difference in annual means --- were initially suggested in 1997 with limited knowledge of actual conditions (lack of data). Now that we have several years of monitoring data available, should we consider adjustments to these criteria? [Rationale to tighten SA criteria.]
- 3 simple evaluations were conducted:

1. Benchmark typical within-area correlation (of daily PM2.5 concentrations). [If SA requires a minimum of .6 correlation, but .6 is only average or worse, shouldn't areas/sites need to show better (higher R) to be permitted to use SA?]
2. Compared annual correlations to seasonal correlations. [If there is significant differences between annual and seasonal correlations, shouldn't the minimum criterion be applied on a seasonal basis?]
3. Benchmarked average percent difference in annual site means to annual spatial means. [If SA requires a maximum of 20% difference in annual means (site vs. spatial) but 20% is only average or worse, shouldn't areas/sites need to show better (lower % difference) to be permitted to use SA?]

Is the Tightening of SA Criteria Appropriate?

1. Benchmark typical within-area correlation (of daily PM2.5 concentrations)
   - Procedure:
     - Utilized SP PM2.5 database (11+samples, all 12 quarters ’01-’03).
     - Calculated correlation between all site pairs in each area (CSA or CBSA)
     - Calculated univariate statistics for site correlations at national level
     - Also averaged correlation to area level then calculated univariate statistics for area averages at national level
     - Reran using only sites pairs where DV’s were within 20% tolerance

All Site Pairs

	Site Stats.	Area Average Stats.
N	2227	129
Maximum	0.9899	0.9899
95th	0.9701	0.9732
75th	0.9343	0.9473
Median	0.8993	0.8999
Mean	0.8764	0.8609
25th	0.8521	0.8228
5th	0.7353	0.6019
Minimum	-0.0854	0.3669

Site Pairs Where DV w/in 20%

	Site Stats.	Area Average Stats.
N	1914	122
Maximum	0.9899	0.9899
95th	0.9712	0.9494
75th	0.9397	0.9494
Median	0.9055	0.9044
Mean	0.8942	0.885
25th	0.8618	0.8462
5th	0.7785	0.7172
Minimum	0.3569	0.5217

•More than 95% of all site pairs have a correlation greater then .7
•The median site correlation is about .9
•More than 95% of all areas have an average correlation greater than .6
•The median area average correlation is about .9

Is the Tightening of SA Criteria Appropriate?

2. Compared annual correlations to seasonal correlations.
   - Procedure:
     - •Utilized SP PM2.5 database (11+samples, all 12 quarters ’01-’03).
     - •Calculated correlation between all site pairs in each area (CSA or CBSA)
       - •Calculated correlation for all paired data points (‘annual’)
       - •Calculated correlation for all paired data points by aggregate quarter (e.g., ‘Q1’= all pairs in 2001-Q1, 2002-Q1, and 2003- Q1) [‘Seasonal’]

•Of the 2227 site pairs:

- **•There was an average difference of about 13% between the annual correlation and the minimum seasonal correlation.**
- **•The median difference is about 6%.**
- **•More than 25% of the pairs had a difference of more than .11 R**
- **•In about 8% of the situations where the ‘annual’ R was > .6, the minimum seasonal R was < .6.**

Is the Tightening of SA Criteria Appropriate?

2. Compared annual correlations to seasonal correlations, cont.

Area	Site 1	Site 2	Correlations						
			Aggregate	Q1	Q2	Q3	Q4	Minimum Quarterl Correlatio n	Difference (Annual - Min_Q)
CSA_Salt Lake City-O	490353007	490030003	0.869	0.931	0.829	0.218	0.917	0.218	0.650
CSA_Seattle-Tacoma-(	530670013	530330037	0.740	0.793	0.863	0.156	0.755	0.156	0.584
CSA_Seattle-Tacoma-(	530670013	530330057	0.750	0.716	0.868	0.185	0.763	0.185	0.564
CSA_Las Vegas-Paradi	320031019	320030022	0.843	0.309	0.856	0.862	0.813	0.309	0.535
CSA_Seattle-Tacoma-(	530611007	530330037	0.807	0.778	0.905	0.315	0.833	0.315	0.492
CSA_Salt Lake City-O	490571003	490353007	0.865	0.921	0.581	0.379	0.909	0.379	0.486
CBSA_Portland-Vancou	410671003	410090004	0.762	0.743	0.769	0.288	0.838	0.288	0.474
CSA_Salt Lake City-O	490350003	490030003	0.912	0.932	0.811	0.445	0.919	0.445	0.466
CSA_Atlanta-Sandy Sp	130670003	130630091	0.800	0.346	0.842	0.869	0.914	0.346	0.454
CSA_Atlanta-Sandy Sp	132230003	130670003	0.816	0.372	0.877	0.912	0.905	0.372	0.444
CSA_Seattle-Tacoma-(	530611007	530330057	0.786	0.769	0.814	0.346	0.780	0.346	0.440
CBSA_Provo-Orem, UT	490495010	490494001	0.922	0.937	0.861	0.510	0.970	0.510	0.412
CSA_Atlanta-Sandy Sp	131210032	130670003	0.862	0.459	0.945	0.963	0.930	0.459	0.403
CSA_Seattle-Tacoma-(	530610005	530330057	0.874	0.857	0.863	0.474	0.883	0.474	0.400
CSA_Atlanta-Sandy Sp	131390003	130670003	0.774	0.381	0.843	0.892	0.782	0.381	0.393
CSA_Atlanta-Sandy Sp	131210039	130670003	0.703	0.317	0.823	0.900	0.627	0.317	0.386
CSA_Seattle-Tacoma-(	530670013	530330080	0.648	0.654	0.852	0.264	0.694	0.264	0.384
CSA_San Juan-Caguas	720610005	720530003	0.707	0.324	0.830	0.848	0.671	0.324	0.383
CBSA_Provo-Orem, UT	490495010	490490002	0.946	0.959	0.872	0.564	0.983	0.564	0.381
CBSA_Portland-Vancou	410510246	410090004	0.741	0.752	0.873	0.361	0.879	0.361	0.380
CSA_Omaha-Council B	310550052	310250002	0.739	0.870	0.938	0.918	0.359	0.359	0.380
CBSA_Tucson, AZ	040191028	040190011	0.793	0.419	0.892	0.829	0.924	0.419	0.375
CSA_Omaha-Council B	311530007	310250002	0.744	0.950	0.950	0.854	0.372	0.372	0.372
CSA_Salt Lake City-O	490570007	490353006	0.936	0.943	0.877	0.564	0.952	0.564	0.372
CSA_Salt Lake City-O	490571003	490350003	0.909	0.914	0.541	0.550	0.929	0.541	0.367
CSA_New York-Newark	340273001	090011123	0.787	0.420	0.863	0.923	0.780	0.420	0.367
CSA_Atlanta-Sandy Sp	131210039	130630091	0.764	0.398	0.705	0.881	0.856	0.398	0.365
CSA_Omaha-Council B	310550019	310250002	0.775	0.909	0.956	0.946	0.416	0.416	0.358
CSA_New York-Newark	340392003	340210008	0.876	0.521	0.903	0.967	0.904	0.521	0.355
CSA_Milwaukee-Racine	550790099	550790010	0.857	0.968	0.502	0.991	0.987	0.502	0.355
CSA_Salt Lake City-O	490571003	490353006	0.912	0.917	0.673	0.561	0.925	0.561	0.351
CSA_Omaha-Council B	310250002	191550009	0.751	0.866	0.901	0.931	0.400	0.400	0.351
CSA_Milwaukee-Racine	550790043	550790010	0.812	0.927	0.461	0.992	0.952	0.461	0.351
CSA_Little Rock-Nort	050690006	050450002	0.793	0.443	0.664	0.855	0.857	0.443	0.350
CSA_Atlanta-Sandy Sp	132230003	131210039	0.706	0.356	0.747	0.868	0.561	0.356	0.350
CSA_Washington-Balti	240030019	110010043	0.870	0.521	0.973	0.830	0.892	0.521	0.349
CBSA_Pocatello, ID	160770011	160050015	0.755	0.785	0.412	0.720	0.763	0.412	0.343
CSA_Salt Lake City-O	490571003	490570007	0.955	0.977	0.613	0.679	0.956	0.613	0.342
CBSA_Honolulu, HI	150031001	150030010	0.436	0.790	0.581	0.722	0.095	0.095	0.341
CSA_Washington-Balti	511071005	110010043	0.850	0.509	0.965	0.809	0.847	0.509	0.341
CBSA_Albuquerque, NI	350439004	350010024	0.606	0.267	0.730	0.814	0.696	0.267	0.339
CSA_Oklahoma City-SI	401091037	400819005	0.801	0.467	0.814	0.947	0.826	0.467	0.334
CSA_Salt Lake City-O	490571003	490350012	0.915	0.948	0.582	0.623	0.894	0.582	0.333
CSA_Salt Lake City-O	490570007	490030003	0.947	0.970	0.835	0.616	0.941	0.616	0.331
CSA_Washington-Balti	240251001	110010043	0.819	0.491	0.936	0.750	0.863	0.491	0.328
CSA_Salt Lake City-O	490353006	490030003	0.921	0.934	0.788	0.595	0.927	0.595	0.327
CSA_New York-Newark	340270004	340171003	0.843	0.520	0.924	0.942	0.908	0.520	0.323
CSA_Washington-Balti	245100007	110010043	0.847	0.526	0.900	0.846	0.878	0.526	0.322
CSA_Washington-Balti	510130020	110010043	0.904	0.587	0.981	0.890	0.948	0.587	0.318
CSA_San Juan-Caguas	720690001	720610005	0.670	0.353	0.873	0.900	0.717	0.353	0.317
CSA_Omaha-Council B	310550051	310250002	0.741	0.935	0.820	0.958	0.424	0.424	0.317
CSA_Seattle-Tacoma-(	530610005	530330037	0.828	0.807	0.920	0.511	0.825	0.511	0.316
CSA_New York-Newark	340270004	340230006	0.893	0.577	0.923	0.948	0.928	0.577	0.316
CSA_New York-Newark	340273001	340270004	0.931	0.616	0.961	0.978	0.927	0.616	0.315
CSA_Little Rock-Nort	051191004	050450002	0.809	0.494	0.665	0.894	0.818	0.494	0.315
CSA_Oklahoma City-SI	401090035	400819005	0.816	0.502	0.840	0.967	0.794	0.502	0.314
CSA_Washington-Balti	240313001	110010043	0.862	0.549	0.977	0.816	0.880	0.549	0.313

- **The table on the left shows examples of where there are large differences between the ‘annual’ correlations and the ‘seasonal’ correlations**
- **There are instances where the ‘annual’ correlation is more than 4 times the minimum ‘seasonal’ correlation.**
- **In most of these extreme cases, the ‘annual’ still meets the current suggested minimum of .6**

Is the Tightening of SA Criteria Appropriate?

3. Benchmarked average percent difference in annual site means versus annual spatial means)
   - Procedure:
     - Utilized SP PM2.5 database (11+samples, all 12 quarters ’01-’03).
     - Calculated average difference between annual site mean and annual area spatial mean. Note that all complete sites in the area were included in the analyses even though this would often not be the case in ‘real world’ (since there are many situations where real low sites would not be included based on correlation, etc.) Thus, the differences shown below are biased high.

Average percent difference in annual site mean versus annual spatial mean

N	1722
Maximum	151.5%
95th	23.7%
75th	9.8%
Median	5.0%
Mean	8.1%
25th	2.0%
5th	0.4%
Minimum	0.0%

- **The median (absolute) difference is 5%**
- **The average difference is 8%**
- **In less than 25% of all cases is the difference greater than 10%**
- **The current SA criterion of 20% is between the 90^{th} and 95^{th} percentile.**

What would tightened criteria yield?

Using criteria of .9 seasonal correlation and +/-10 % difference in annual means. Using annual std level of 15.0				Using criteria of .9 seasonal correlation and +/-10 % difference in annual means. Using annual std level of 14.0			
		Could use spatial averaging	Could use SA to meet 15.0 annual std			Could use spatial averaging	Could use SA to meet 14.0 annual std
Number of areas		12	2	Number of areas		18	1
Total population		22,327,531	1,233,836	Total population		27,499,635	1,381,287
Area distribution statistics:				Area distribution statistics:			
Difference in area DV's (ug/m3)	mean	0.52	0.50	Difference in area DV's (ug/m3)	mean	0.44	0.70
	max	1.2	0.7		max	1.2	0.7
	p95	1.2	0.7		p95	1.2	0.7
	p75	0.8	0.7		p75	0.7	0.7
	med	0.4	0.5		med	0.3	0.7
	p25	0.2	0.3		p25	0.2	0.7
	p05	0.1	0.3		p05	0.0	0.7
	min	0.1	0.3		min	0.0	0.7

- By tightening the annual mean difference criterion and the correlation criterion, much fewer areas would qualify for SA. Using a .9 quarterly correlation cutoff (as shown above) would narrow the option to 18 or fewer areas. The average difference in area means (SA versus regular) would also decline to about .4-.5 ug/m3. Total population for these areas is 22-27 million.
- Only 1 or 2 of these areas could use SA to meet the annual std NAAQS with their spatial average when they couldn't with their regular site-based DV. The realized reduction in DV for these areas would be .5-.7 ug/m3. Only 1 million people live in those areas.

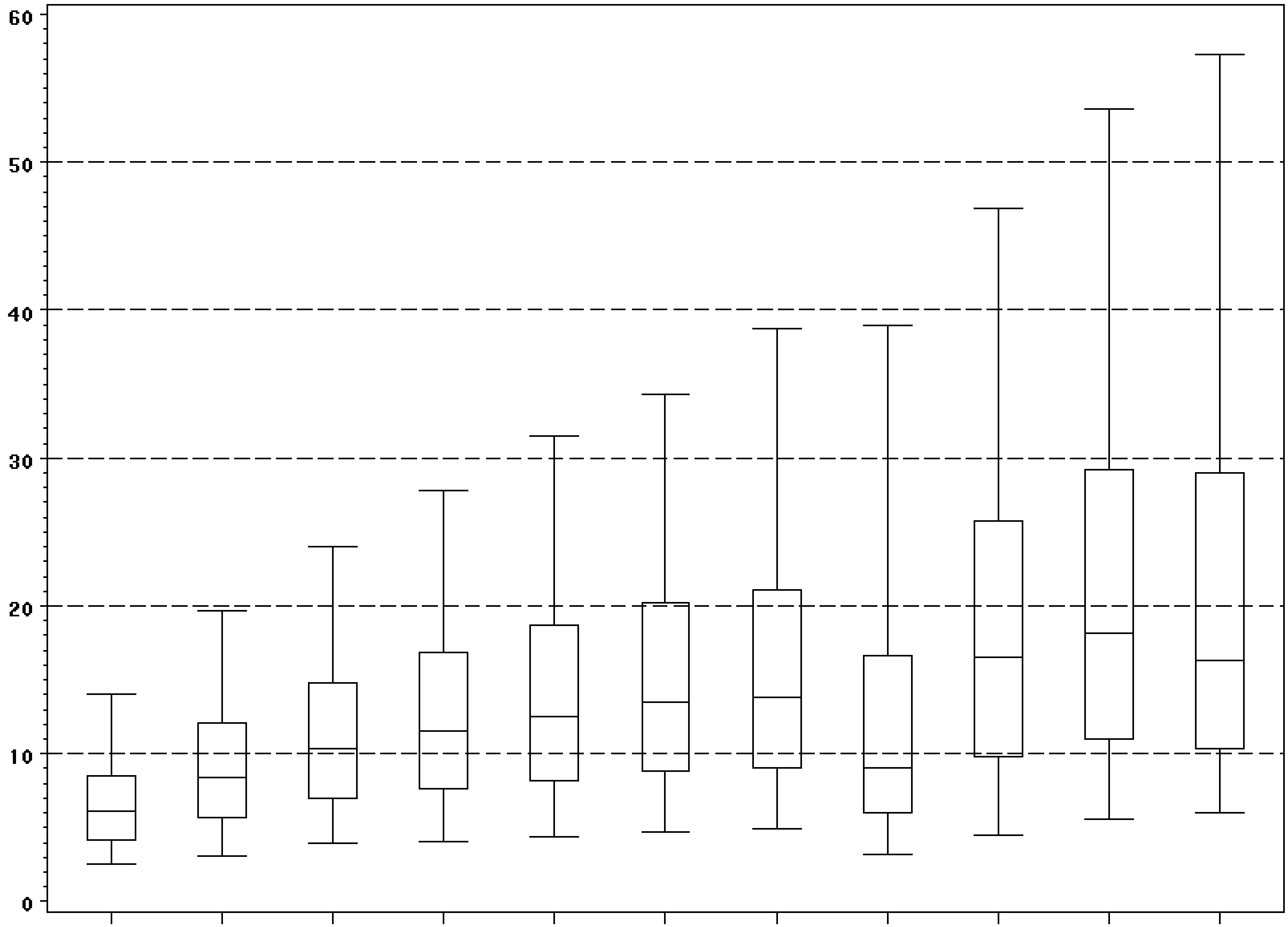
PM2.5 Evaluation of High Concentrations

- Purpose:
  - To identify the minimum number of days permitted per year to exceed the annual 98th, 99th, etc. percentiles.
  - To evaluate the (entire) daily distributions of data plotted by 98th (and 99th) percentile level intervals.
  - To evaluate the daily distributions of data exceeding site-level 98th (and 99th) DV's plotted by 98th (and 99th) percentile intervals.
  - To ascertain the actual number and percentage of days (site average, minimum, & maximum), for the 3-year period 2001-2003, where the concentration was significantly above the site 98th or 99th percentiles. [Significant defined as 5+ ug/m3.]

Number of Exempted Days Per Year for Percentile Metrics

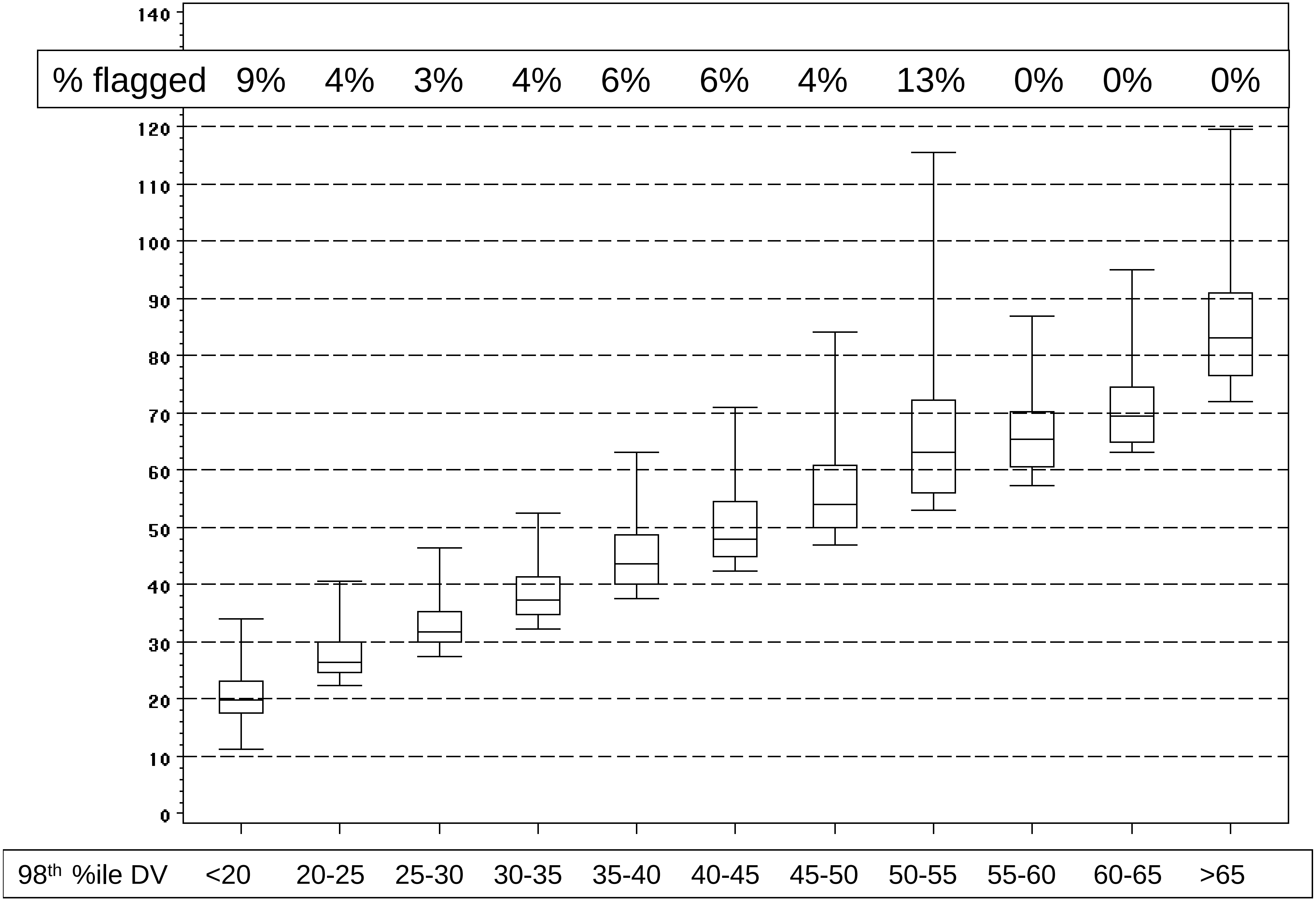
Sampling Frequency	Number of Sample Days	Number of Sample Days Above Stated Percentile				
		95th	96th	97th	98th	99th
Every Day	~ 365	18	14	10	7	3
Every 3rd Day	~ 122	6	4	3	2	1
Every 6th Day	~ 61	3	2	1	1	0

Distribution of PM2.5 concentrations by 98th percentile DV interval



98th %ile DV	<20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	>65

Distribution of PM2.5 concentrations > 98th percentile DV, by 98th percentile DV interval



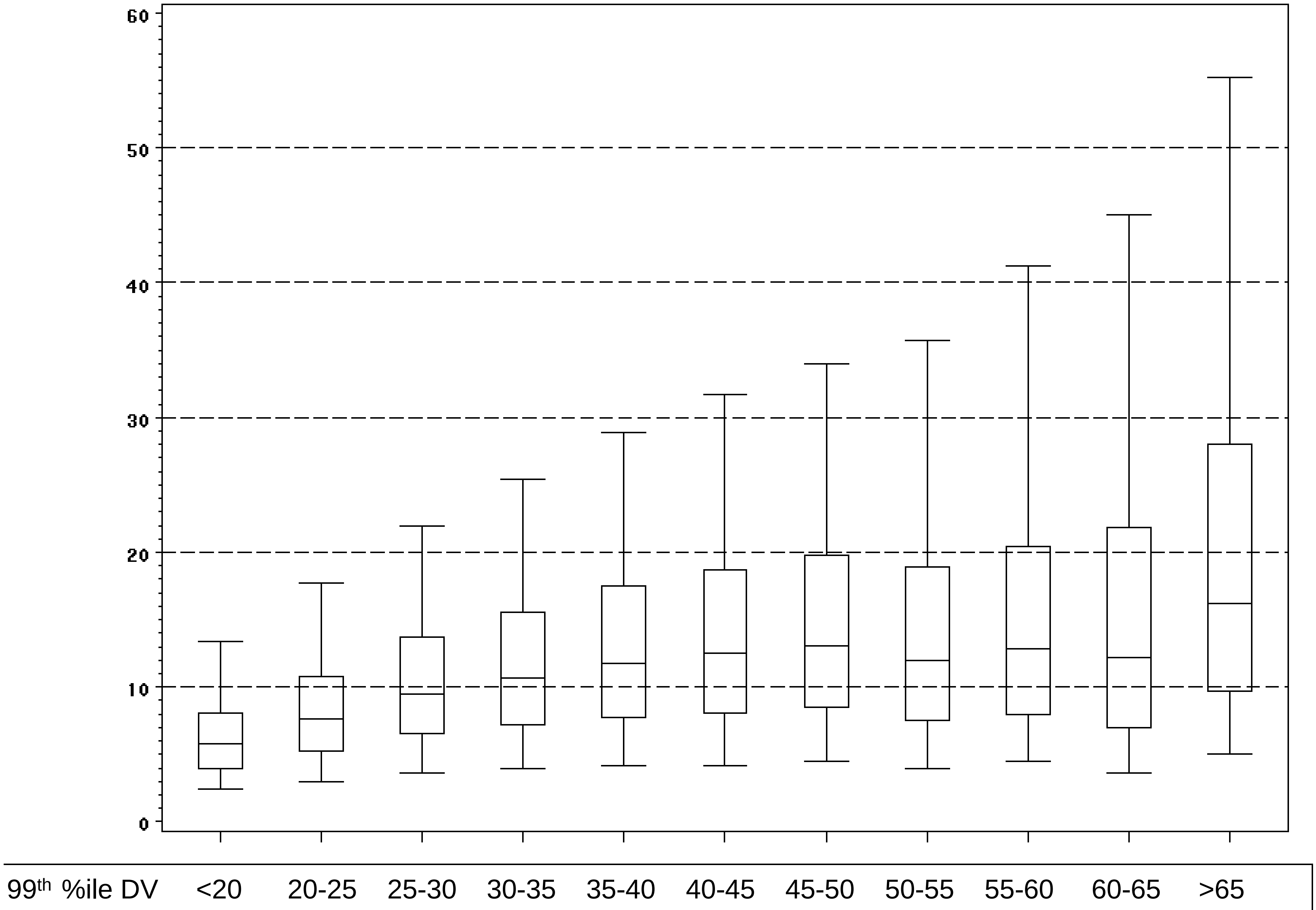
Percent and number of days PM2.5 concentrations exceeded the site 98th percentile DV by more than 5ug/m3, 2001-2003

		Days ≤ P98DV + 5 (but ≥ P98DV)						Days > P98DV + 5					
P98 DV	# sites	Minimum # Days	Mean # Days	Maximum # Days	Minimum % Days	Mean % Days	Maximum % Days	Minimum # Days	Mean # Days	Maximum # Days	Minimum % Days	Mean % Days	Maximum % Days
all	827	0	5.7	24	0.0%	1.3%	4.6%	0	4.1	20	0.0%	0.9%	2.8%
<20	81	0	6.3	21	0.0%	1.6%	3.7%	0	2.2	7	0.0%	0.6%	1.8%
20-25	81	0	7.1	22	0.0%	1.6%	3.4%	0	3.0	11	0.0%	0.7%	1.7%
25-30	192	1	5.6	24	0.3%	1.4%	4.6%	0	3.2	12	0.0%	0.8%	2.8%
30-35	197	0	5.7	19	0.0%	1.3%	3.6%	0	4.0	15	0.0%	0.9%	2.5%
35-40	179	0	5.4	18	0.0%	1.1%	2.3%	1	5.1	16	0.3%	1.1%	2.2%
40-45	53	0	5.7	16	0.0%	1.0%	2.0%	0	6.0	19	0.0%	1.1%	2.8%
45-50	12	0	4.1	13	0.0%	0.7%	1.3%	3	5.5	13	0.4%	1.3%	2.8%
50-55	11	0	4.0	9	0.0%	0.7%	1.7%	2	8.3	20	0.9%	1.4%	2.1%
55-60	7	0	4.9	9	0.0%	0.7%	1.2%	4	7.3	12	1.0%	1.3%	1.7%
60-65	7	3	4.9	7	0.7%	1.3%	2.1%	3	6.0	12	0.9%	1.5%	2.6%
>65	7	1	2.7	5	0.3%	0.4%	0.6%	4	9.9	18	1.2%	1.9%	2.8%

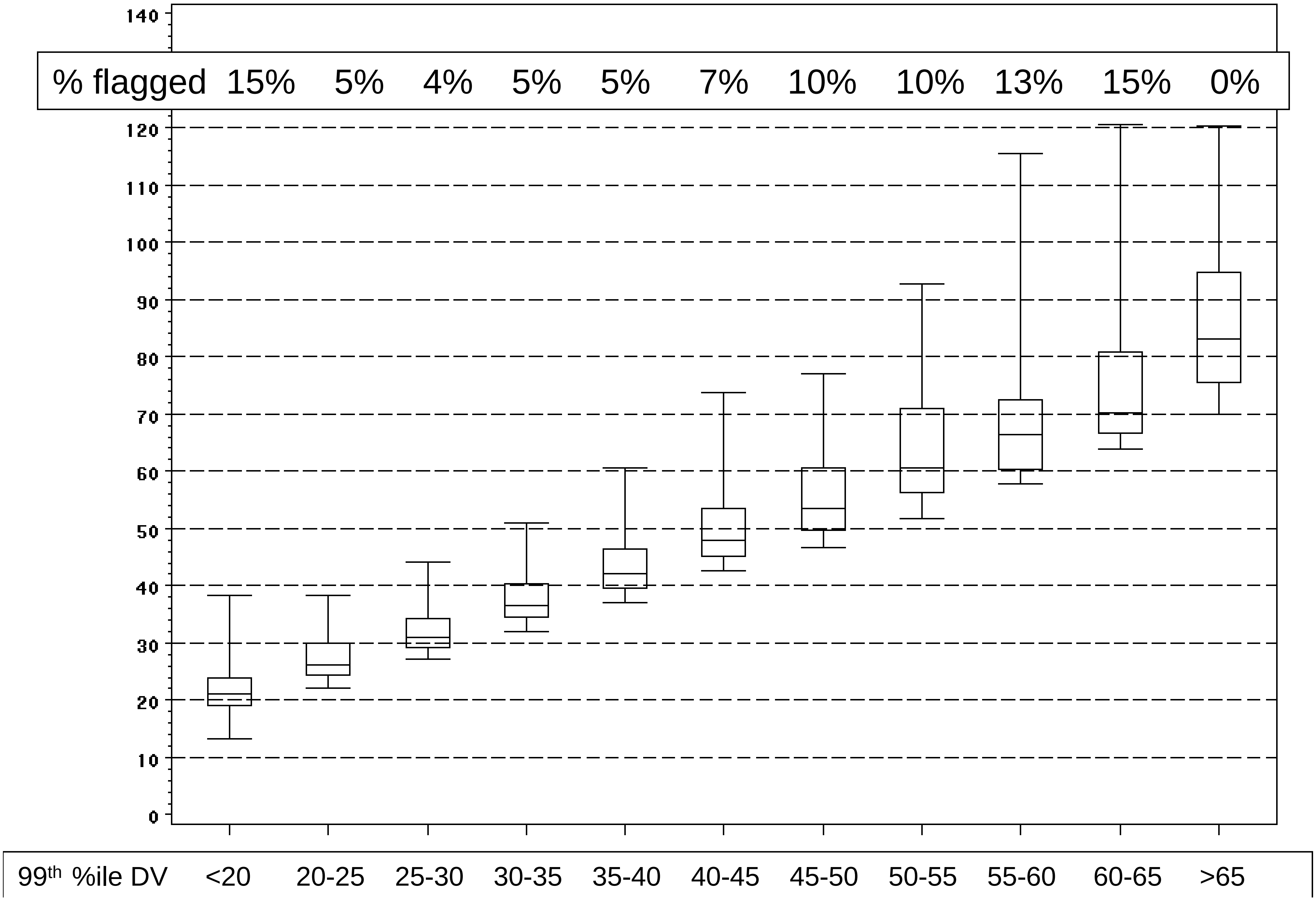
Maximum number of days in any one year (2001-2003) that a site exceeded it's 3-year 98th or 99th percentile DV.

- Site 410290133 exceeded its 98th percentile DV (of 37ug/m3) 20 times in 2002.
- Site 410350004 exceeded its 99th percentile DV (of 65ug/m3) 13 times in 2002.
- The theoretical answer for both is 365 (or 365 for leap-year)!

Distribution of PM2.5 concentrations by 99th percentile 'DV' interval



Distribution of PM2.5 concentrations > 99th percentile 'DV', by 99th percentile DV interval



Percent and number of days PM2.5 concentrations exceeded the site 99th percentile 'DV' by more than 5ug/m3, 2001-2003

		Days ≤ P99DV + 5 (but ≥ P99DV)						Days > P99DV + 5					
P99 DV	# sites	Minimum # Days	Mean # Days	Maximum # Days	Minimum % Days	Mean % Days	Maximum % Days	Minimum # Days	Mean # Days	Maximum # Days	Minimum % Days	Mean % Days	Maximum % Days
all	827	0	3.4	14	0.0%	0.8%	3.9%	0	2.3	12	0.0%	0.5%	2.4%
<20	53	0	4.0	12	0.0%	1.0%	2.4%	0	1.4	6	0.0%	0.3%	0.9%
20-25	58	0	3.5	11	0.0%	0.9%	2.8%	0	1.4	6	0.0%	0.4%	1.0%
25-30	121	0	4.1	11	0.0%	1.1%	2.6%	0	1.6	7	0.0%	0.4%	2.4%
30-35	183	0	3.6	13	0.0%	0.9%	3.9%	0	1.9	7	0.0%	0.4%	1.1%
35-40	161	0	3.3	12	0.0%	0.8%	1.8%	0	2.3	8	0.0%	0.5%	1.4%
40-45	145	0	3.2	14	0.0%	0.6%	1.9%	0	3.0	8	0.0%	0.6%	2.1%
45-50	53	0	2.9	10	0.0%	0.5%	1.3%	0	3.4	7	0.0%	0.6%	1.4%
50-55	19	0	2.1	5	0.0%	0.4%	1.1%	1	3.7	12	0.3%	0.7%	1.5%
55-60	7	0	1.9	5	0.0%	0.4%	0.9%	2	3.6	7	0.4%	0.8%	1.1%
60-65	13	0	3.2	7	0.0%	0.6%	1.3%	1	4.0	10	0.3%	0.8%	1.3%
>65	14	0	2.1	6	0.0%	0.4%	0.9%	2	5.5	10	0.6%	1.2%	2.3%

Predicted Percentage of Counties w/ Monitors Not Likely to Meet Alternative PM Standards

PM2.5, PM10-2.5, and PM10

Description of Databases

- PM2.5
  - Utilized 2001-2003 production design value database (12 quarters, every site) with two implemented exceptions: 1) Sites were deemed complete if they had 11+ samples per quarter irregardless of DV level. [In 'official DV's, 11+ per quarter is only sufficient if the DV exceeds the level of the std.]; and 2) Data 'substitution' was not employed. [In official DV's, incomplete sites were run through several tests to see if they could be deemed complete even if they don't meet the precise CFR requirements; these test include substitution of high / low values].
  - **Excluded concurred events (as in 'official' DV's).**
  - 827 sites. 562 counties. Less counties than last year (693) with '12,8,4' approach but, 1) better data quality, 2) site differences reflect spatial not temporal differences…. lots of PM2.5 concentration declines between '01 and '03; and 3) not much difference in total populations
- PM10-2.5
  - Utilized 'regular' characterization database and 'extended' db. Both utilized '12,8,4' approach since requiring 12Q's would significantly reduce db size. **Excluded sites deemed source oriented and <u>not</u> population exposure!** Excluded event-flagged data.
  - Regular: Same logic as last year: Collocated FRM PM10 and FRM PM2.5 (and small amount of dichot.) When multiple data pairs exist for same site-day, PMc was differenced from all pairs and averaged.] 487 sites, 350 counties.
  - Extended: Includes 'regular' db plus data pairs from non-collocated (but nearby) sites; PMc anchored at PM10 site. Assumption: PM2.5 fairly homogenous, PM10 not. Why include non-collocated? Many 'high' PM10 sites do not have collocated PM2.5 because of disparate monitoring objectives (highest conc. vs. population exposure.). Hence, by not doing, we would be ignoring many potentially high PMc locations. 712 sites. 382 counties.
- PM10
  - Utilized production DV database. Required 75% data capture. **Excluded concurred events.**
  - 585 counties.

Estimated Number/Population/Percentage of Counties Violating PM2.5 Alternative NAAQS, Annual Only & Combination Annual / 98th Percentile

Alternative Standards and Levels (□g/m^3)	Population in monitored counties not likely to meet stated standard and level (1000's)	Percent population (county based) not likely to meet stated standard and level	Number of counties not likely to meet stated standard and level	Percent number of counties not likely to meet stated standard and level*								
	U.S. Total				Northeast	Southeast	Industrial Midwest	Upper Midwest	Southwest	Northwest	Southern CA	Outside Regions**
Annual only												
15	55,855	30%	78	14%	19%	7%	29%	0%	0%	4%	60%	0%
14	76,934	41%	140	25%	28%	21%	51%	0%	5%	5%	67%	0%
13	102,444	55%	224	40%	47%	40%	76%	4%	5%	7%	67%	0%
12	122,454	66%	304	54%	70%	61%	89%	12%	5%	12%	67%	0%
Combined Annual / 24-hour												
15 / 65	55,855	30%	78	14%	19%	7%	29%	0%	0%	4%	60%	0%
15 / 50	58,391	31%	82	15%	19%	7%	29%	0%	0%	9%	60%	0%
15 / 45	60,757	33%	87	15%	19%	7%	29%	0%	10%	12%	60%	0%
15 / 40	65,296	35%	94	17%	20%	7%	30%	0%	10%	19%	60%	0%
15 / 35	89,779	48%	153	27%	45%	8%	47%	0%	10%	36%	60%	7%
15 / 30	133,216	72%	289	51%	78%	29%	87%	6%	19%	51%	80%	13%
15 / 25	159,187	86%	441	78%	98%	77%	99%	51%	43%	65%	80%	13%
14 / 65	76,934	41%	140	25%	28%	21%	51%	0%	5%	5%	67%	0%
14 / 50	79,470	43%	144	26%	28%	21%	51%	0%	5%	10%	67%	0%
14 / 45	81,129	44%	147	26%	28%	21%	51%	0%	10%	12%	67%	0%
14 / 40	84,919	46%	153	27%	28%	21%	52%	0%	10%	19%	67%	0%
14 / 35	101,327	55%	191	34%	45%	22%	58%	0%	10%	36%	67%	7%
14 / 30	134,420	72%	296	53%	78%	33%	88%	6%	19%	51%	80%	13%
15 / 25	159,187	86%	441	78%	98%	77%	99%	51%	43%	65%	80%	13%
13 / 65	102,444	55%	224	40%	47%	40%	76%	4%	5%	7%	67%	0%
13 / 50	103,759	56%	226	40%	47%	40%	76%	4%	5%	10%	67%	0%
13 / 45	105,418	57%	229	41%	47%	40%	76%	4%	10%	12%	67%	0%
13 / 40	108,257	58%	234	42%	47%	40%	76%	4%	10%	19%	67%	0%
13 / 35	115,814	62%	255	45%	53%	40%	77%	4%	10%	36%	67%	7%
13 / 30	137,807	74%	318	57%	78%	43%	90%	8%	19%	51%	80%	13%
13 / 25	159,187	86%	441	78%	98%	77%	99%	51%	43%	65%	80%	13%
12 / 65	122,454	66%	304	54%	70%	61%	89%	12%	5%	12%	67%	0%
12 / 50	122,454	66%	304	54%	70%	61%	89%	12%	5%	12%	67%	0%
12 / 45	123,910	67%	306	54%	70%	61%	89%	12%	10%	14%	67%	0%
12 / 40	126,750	68%	311	55%	70%	61%	89%	12%	10%	20%	67%	0%
12 / 35	132,384	71%	325	58%	70%	61%	89%	12%	10%	36%	67%	7%
12 / 30	144,722	78%	362	64%	84%	62%	94%	14%	19%	51%	80%	13%
12 / 25	159,243	86%	442	79%	98%	78%	99%	51%	43%	65%	80%	13%
Total number of monitored counties (w/ data) ---->				*562*	*83*	*168*	*130*	*49*	*21*	*81*	*15*	*15*
Total population of monitored counties (1000's) ---->				*185,780*	*38,730*	*43,574*	*39,000*	*7,793*	*8,617*	*22,948*	*22,467*	*2,652*

* Based on 2001-2003 data for sites with at least 11 samples per quarter for all 12 quarters. As such, these estimates are not based on the same air quality data that would be used to determine whether an area would attain a given standard or set of standards. These estimates can only approximate the number of counties that are likely not to attain the given standards and should be interpreted with caution.

** "Outside Regions" includes Alaska, Hawaii, Puerto Rico, and the Virgin Islands.

Estimated Number/Population/Percentage of Counties Violating PM2.5 Alternative NAAQS, Annual Only & Combination Annual / 99th Percentile

Alternative Standards and Levels (μg/m³)	Population in monitored counties not likely to meet stated standard and level (1000's)	Percent population (county based) not likely to meet stated standard and level	Number of counties not likely to meet stated standard and level	Percent number of counties not likely to meet stated standard and level*								
	U.S. Total				Northeast	Southeast	Industrial Midwest	Upper Midwest	Southwest	Northwest	Southern CA	Outside Regions**
Annual only												
15	55,855	30%	78	14%	19%	7%	29%	0%	0%	4%	60%	0%
14	76,934	41%	140	25%	28%	21%	51%	0%	5%	5%	67%	0%
13	102,444	55%	224	40%	47%	40%	76%	4%	5%	7%	67%	0%
12	122,454	66%	304	54%	70%	61%	89%	12%	5%	12%	67%	0%
Combined Annual / 24-hour												
15 / 65	55,946	30%	79	14%	19%	7%	29%	0%	0%	5%	60%	0%
15 / 50	61,520	33%	89	16%	19%	7%	29%	0%	10%	15%	60%	0%
15 / 45	65,834	35%	101	18%	24%	7%	32%	0%	10%	21%	60%	0%
15 / 40	86,303	46%	150	27%	47%	9%	42%	0%	10%	36%	67%	7%
15 / 35	126,468	68%	247	44%	72%	17%	77%	0%	19%	51%	80%	13%
15 / 30	151,550	82%	383	68%	96%	54%	97%	35%	38%	59%	80%	13%
15 / 25	165,619	89%	475	85%	100%	86%	99%	69%	48%	73%	87%	13%
14 / 65	77,025	41%	141	25%	28%	21%	51%	0%	5%	6%	67%	0%
14 / 50	81,892	44%	149	27%	28%	21%	51%	0%	10%	15%	67%	0%
14 / 45	84,236	45%	157	28%	30%	21%	52%	0%	10%	21%	67%	0%
14 / 40	99,235	53%	195	35%	48%	23%	57%	0%	10%	36%	73%	7%
14 / 35	129,387	70%	266	47%	72%	27%	78%	0%	19%	51%	80%	13%
14 / 30	151,550	82%	383	68%	96%	54%	97%	35%	38%	59%	80%	13%
15 / 25	165,619	89%	475	85%	100%	86%	99%	69%	48%	73%	87%	13%
13 / 65	102,535	55%	225	40%	47%	40%	76%	4%	5%	9%	67%	0%
13 / 50	106,181	57%	231	41%	47%	40%	76%	4%	10%	15%	67%	0%
13 / 45	108,360	58%	238	42%	49%	40%	76%	4%	10%	21%	67%	0%
13 / 40	116,019	62%	262	47%	59%	40%	77%	4%	10%	36%	73%	7%
13 / 35	135,204	73%	302	54%	75%	40%	85%	4%	19%	51%	80%	13%
13 / 30	152,684	82%	391	70%	96%	58%	97%	35%	38%	59%	80%	13%
13 / 25	165,619	89%	475	85%	100%	86%	99%	69%	48%	73%	87%	13%
12 / 65	122,454	66%	304	54%	70%	61%	89%	12%	5%	12%	67%	0%
12 / 50	124,673	67%	308	55%	70%	61%	89%	12%	10%	16%	67%	0%
12 / 45	126,634	68%	314	56%	71%	61%	89%	12%	10%	22%	67%	0%
12 / 40	132,537	71%	331	59%	75%	62%	89%	12%	10%	36%	73%	7%
12 / 35	143,294	77%	354	63%	80%	62%	92%	12%	19%	51%	80%	13%
12 / 30	154,844	83%	409	73%	96%	68%	98%	35%	38%	59%	80%	13%
12 / 25	165,619	89%	475	85%	100%	86%	99%	69%	48%	73%	87%	13%
Total number of monitored counties (w/ data) ---->				*562*	*83*	*168*	*130*	*49*	*21*	*81*	*15*	*15*
Total population of monitored counties (1000's) ---->				*185,780*	*38,730*	*43,574*	*39,000*	*7,793*	*8,617*	*22,948*	*22,467*	*2,652*

* Based on 2001-2003 data for sites with at least 11 samples per quarter for all 12 quarters. As such, these estimates are not based on the same air quality data that would be used to determine whether an area would attain a given standard or set of standards. These estimates can only approximate the number of counties that are likely not to attain the given standards and should be interpreted with caution.

** "Outside Regions" includes Alaska, Hawaii, Puerto Rico, and the Virgin Islands.

Note: 'Annual only' data same as preceding slide

Estimated Number/Population/Percentage of Counties Violating PM10-2.5 Alternative NAAQS Levels, 98th Percentile

Alternative Standard Levels ($\mu g/m^3$)	Population in monitored counties not likely to meet stated level (1000's)	Percent population (county based) not likely to meet stated level	Number of counties not likely to meet stated level	Percent number of counties not likely to meet stated level*								
	U.S. Total				Northeast	Southeast	Industrial Midwest	Upper Midwest	Southwest	Northwest	Southern CA	Outside Regions**
100	7,497	5%	11	3%	2%	1%	0%	3%	20%	1%	20%	0%
95	8,314	6%	13	3%	2%	1%	0%	3%	25%	1%	20%	7%
90	9,014	6%	15	4%	2%	1%	0%	3%	30%	1%	27%	7%
85	10,493	7%	19	5%	2%	1%	0%	6%	35%	3%	27%	7%
80	11,435	8%	24	6%	4%	1%	1%	9%	40%	3%	27%	14%
75	13,410	9%	29	8%	4%	2%	3%	12%	40%	6%	27%	14%
70	26,447	18%	38	10%	4%	5%	5%	15%	40%	7%	40%	21%
65	28,633	19%	46	12%	5%	5%	7%	15%	45%	10%	47%	29%
60	36,475	24%	60	16%	5%	5%	14%	24%	55%	13%	47%	43%
55	54,010	36%	72	19%	9%	9%	14%	30%	55%	13%	67%	57%
50	57,432	38%	86	23%	11%	10%	16%	30%	65%	19%	67%	71%
45	66,552	44%	110	29%	14%	17%	18%	42%	70%	28%	73%	79%
40	73,112	49%	136	36%	16%	21%	22%	55%	70%	44%	73%	86%
35	82,785	55%	158	41%	21%	22%	33%	64%	80%	49%	80%	86%
30	100,255	67%	202	53%	33%	33%	45%	70%	80%	66%	87%	93%
25	111,268	74%	243	64%	46%	48%	58%	85%	85%	73%	93%	93%
Total number of monitored counties (w/ data) ---->				*382*	*57*	*82*	*73*	*33*	*20*	*88*	*15*	*14*
Total population of monitored counties (1000's) ---->				*150,595*	*27,529*	*33,988*	*26,988*	*5,837*	*9,436*	*21,842*	*22,467*	*2,509*

* Based on 2001-2003 data for sites with 4, 8, or 12 consecutive quarters with at least 11 samples per quarter. As such, these estimates are not based on the same air quality data that would be used to determine whether an area would attain a given standard or set of standards. These estimates can only approximate the number of counties that are likely not to attain the given standards and should be interpreted with caution.

** "Outside Regions" includes Alaska, Hawaii, Puerto Rico, and the Virgin Islands.

Estimated Number/Population/Percentage of Counties Violating PM10-2.5 Alternative NAAQS Levels, 99th Percentile

Alternative Standard Levels ($\mu g/m^3$)	Population in monitored counties not likely to meet stated level (1000's)	Percent population (county based) not likely to meet stated level	Number of counties not likely to meet stated level	Percent number of counties not likely to meet stated level*								
	U.S. Total				Northeast	Southeast	Industrial Midwest	Upper Midwest	Southwest	Northwest	Southern CA	Outside Regions**
100	19,222	13%	22	6%	4%	2%	0%	3%	40%	1%	40%	14%
95	20,001	13%	27	7%	4%	2%	3%	3%	45%	3%	40%	14%
90	20,934	14%	33	9%	5%	2%	4%	6%	50%	6%	40%	14%
85	29,819	20%	44	12%	5%	4%	7%	12%	55%	11%	40%	14%
80	32,565	22%	48	13%	5%	4%	8%	15%	60%	13%	40%	14%
75	36,037	24%	54	14%	5%	6%	10%	15%	60%	13%	53%	21%
70	39,387	26%	62	16%	9%	9%	10%	21%	60%	14%	60%	21%
65	47,475	32%	79	21%	11%	10%	14%	33%	65%	17%	60%	50%
60	57,922	38%	92	24%	12%	11%	16%	33%	70%	23%	67%	64%
55	65,778	44%	110	29%	12%	12%	18%	48%	70%	33%	73%	71%
50	71,217	47%	129	34%	18%	17%	23%	52%	70%	40%	73%	79%
45	79,471	53%	155	41%	18%	24%	27%	58%	80%	51%	87%	86%
40	84,686	56%	170	45%	21%	24%	34%	70%	80%	55%	87%	93%
35	100,162	67%	204	53%	32%	34%	45%	79%	80%	64%	93%	93%
30	109,160	72%	238	62%	42%	45%	56%	85%	85%	73%	93%	93%
25	123,100	82%	288	75%	56%	66%	68%	94%	90%	85%	100%	93%
Total number of monitored counties (w/ data) ---->				*382*	*57*	*82*	*73*	*33*	*20*	*88*	*15*	*14*
Total population of monitored counties (1000's) ---->				*150,595*	*27,529*	*33,988*	*26,988*	*5,837*	*9,436*	*21,842*	*22,467*	*2,509*

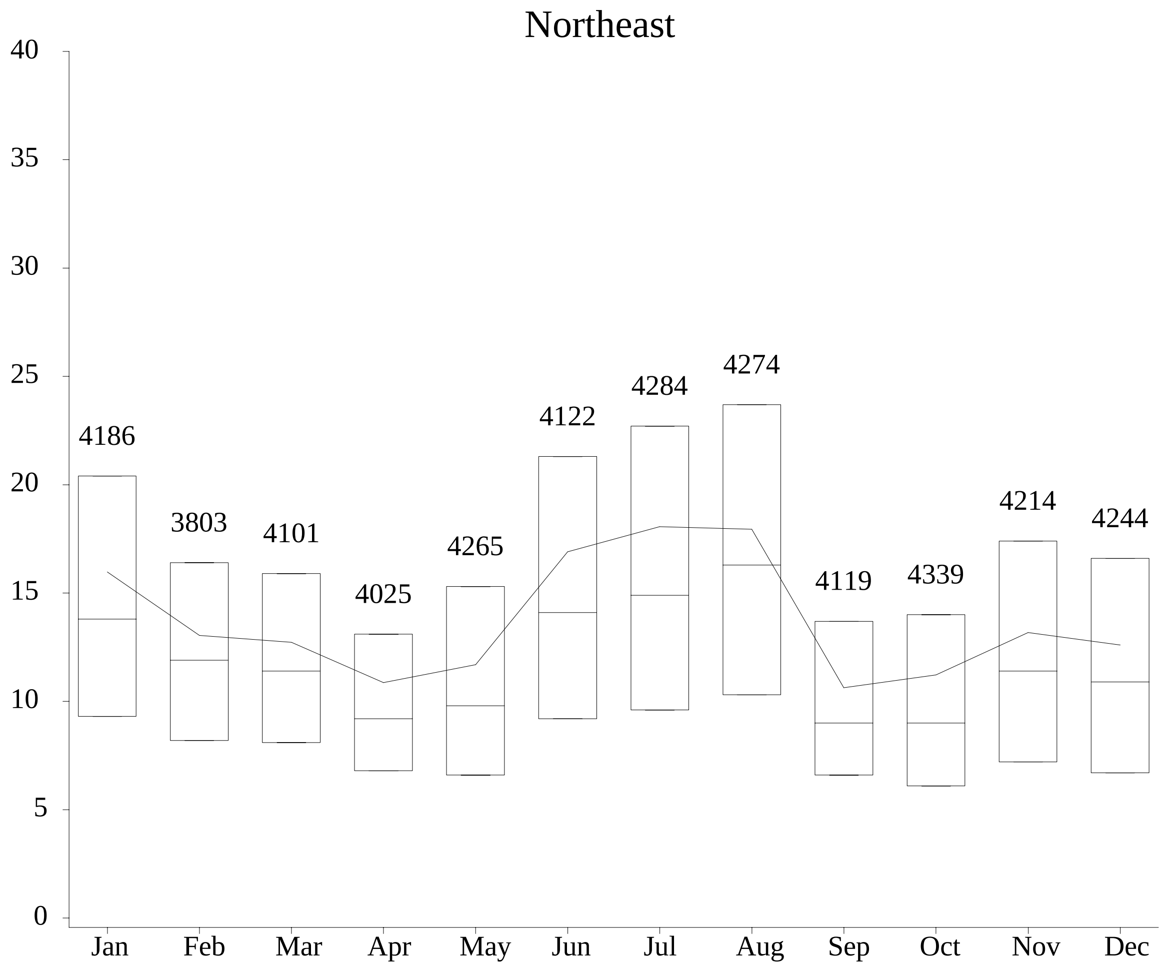
* Based on 2001-2003 data for sites with 4, 8, or 12 consecutive quarters with at least 11 samples per quarter. As such, these estimates are not based on the same air quality data that would be used to determine whether an area would attain a given standard or set of standards. These estimates can only approximate the number of counties that are likely not to attain the given standards and should be interpreted with caution.

** "Outside Regions" includes Alaska, Hawaii, Puerto Rico, and the Virgin Islands.

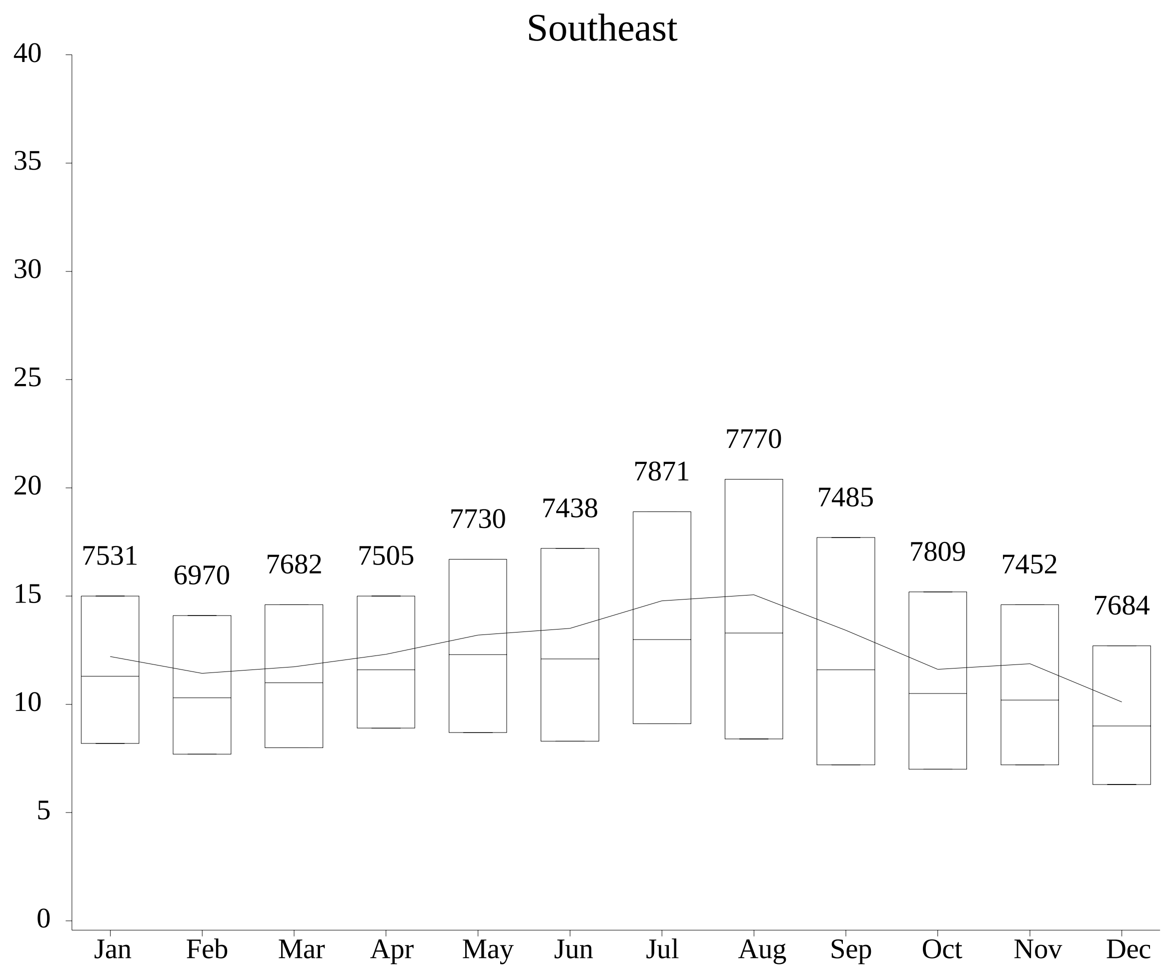
Number/Population/Percentage of Counties Violating PM10 NAAQS

Alternative Standards and Levels	Statistics for counties exceeding levels			
	Number	Percent of Total	Population (1000's)	Percent of Total
Total Number and Population of Monitored Counties				
Total	**585**	**100%**	**170,157**	**100%**
24-Hour Level (Expected Exceedance)				
150	45	8%	18,626	11%
Annual Level				
50	4	1%	3,765	2%
Combined Annual / 24-Hour levels				
50 / 150	45	8%	18,626	495%

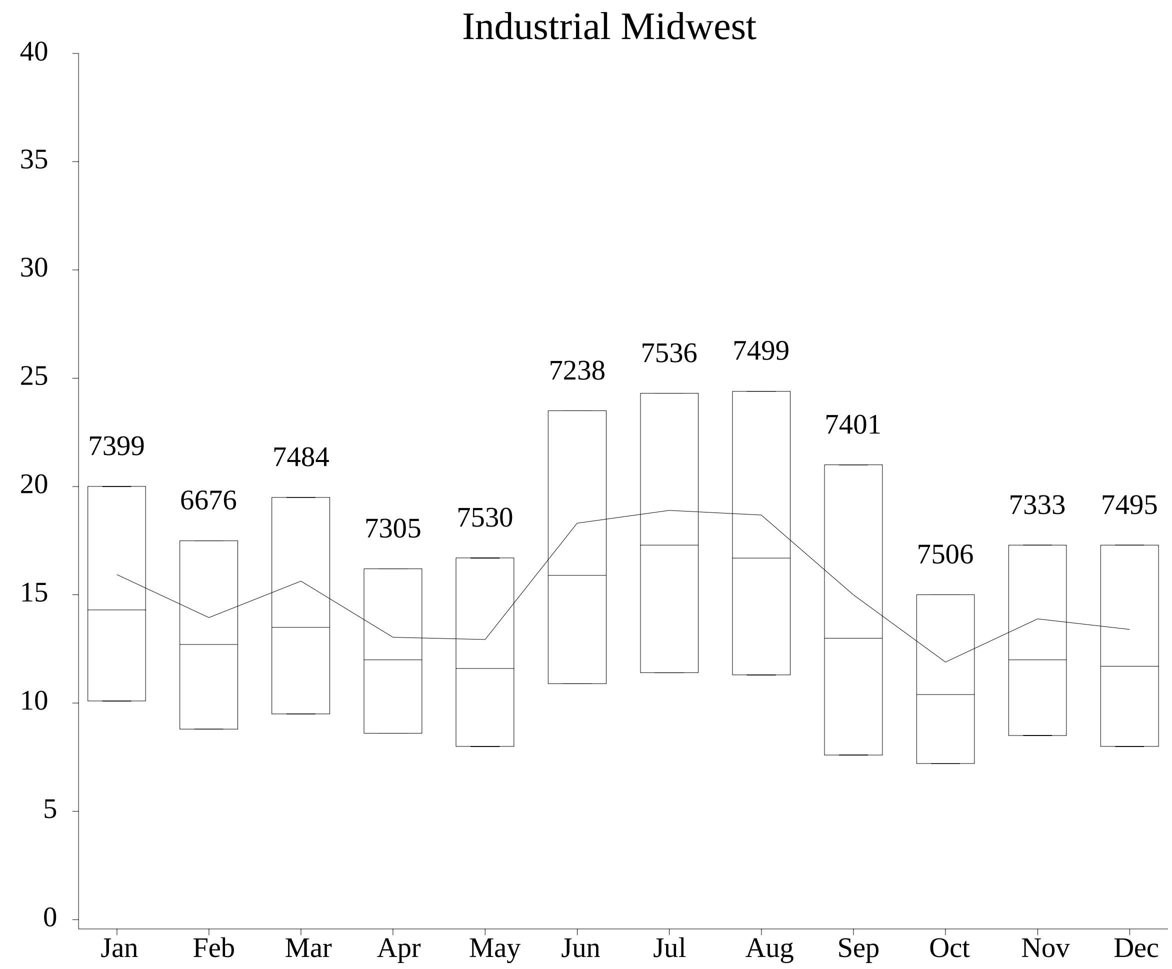
Urban 24-hour average $PM_{2.5}$ concentration distributions by region and month, 2001-2003.



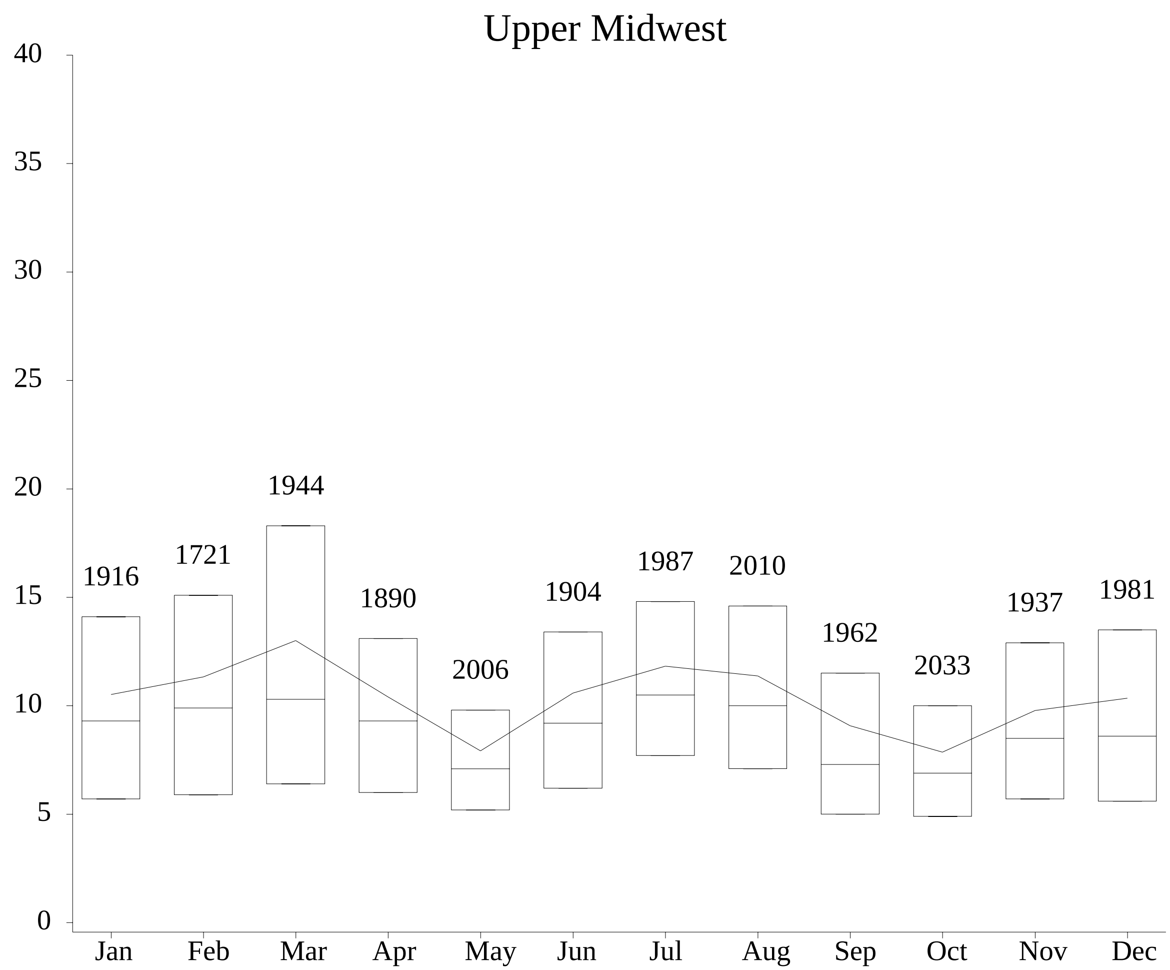
Urban 24-hour average $PM_{2.5}$ concentration distributions by region and month, 2001-2003.



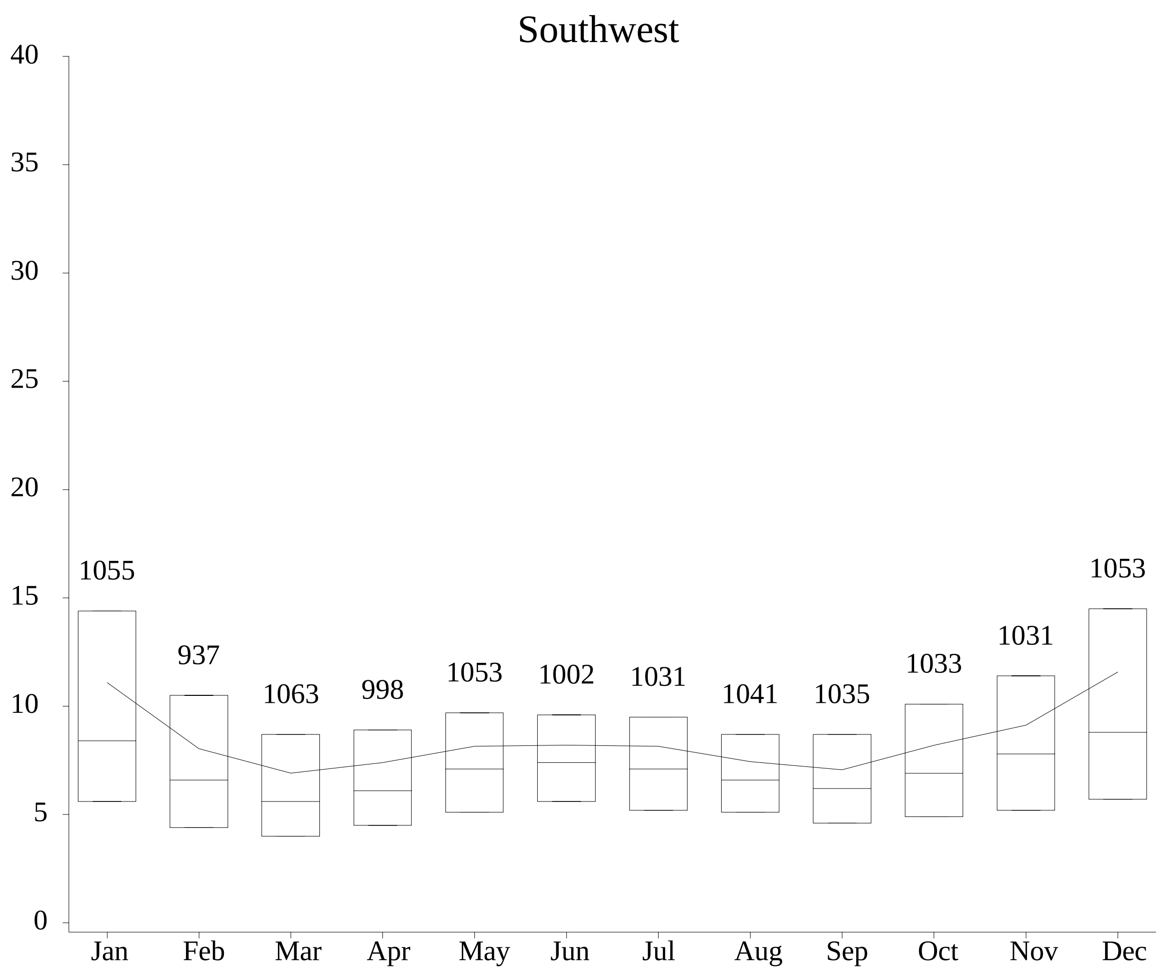
Urban 24-hour average $PM_{2.5}$ concentration distributions by region and month, 2001-2003.



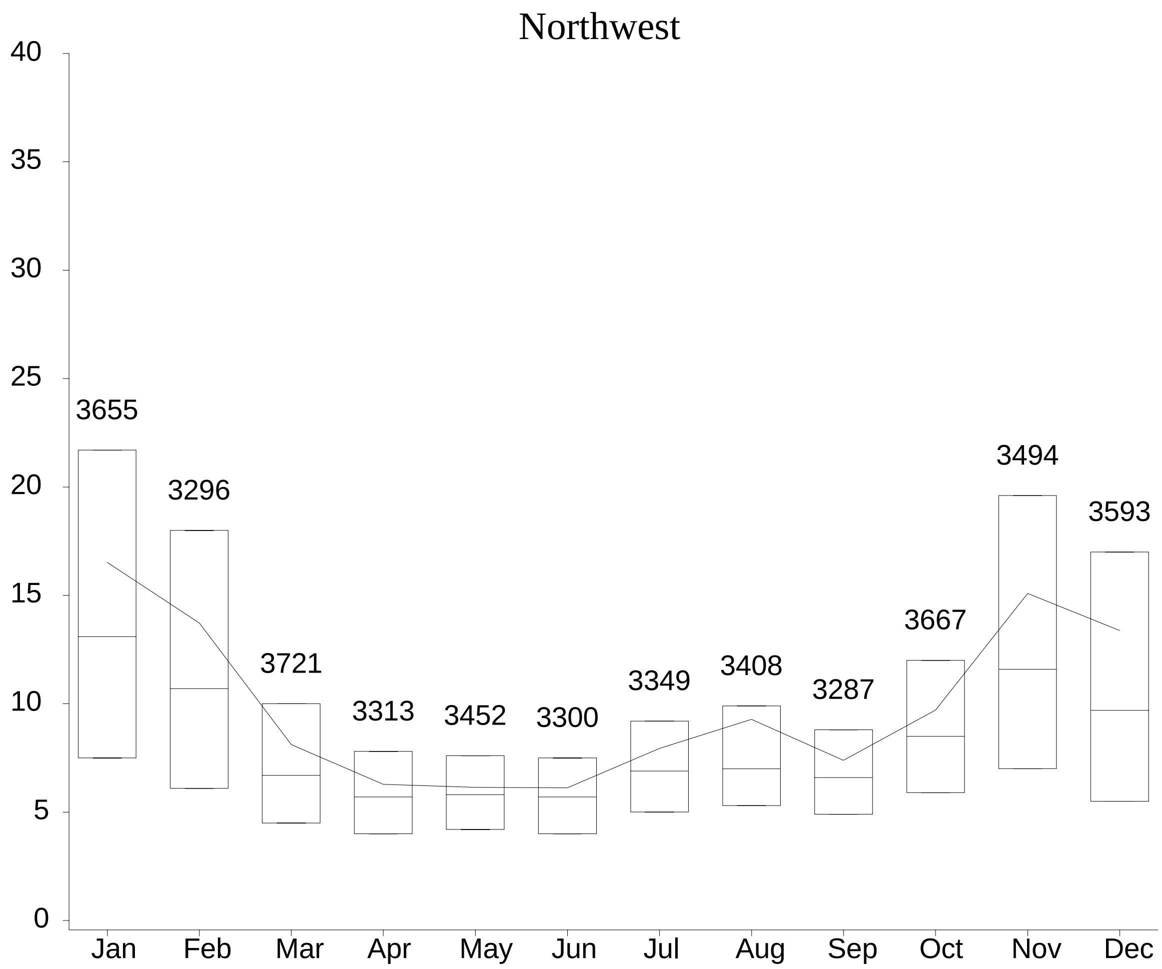
Urban 24-hour average $PM_{2.5}$ concentration distributions by region and month, 2001-2003.



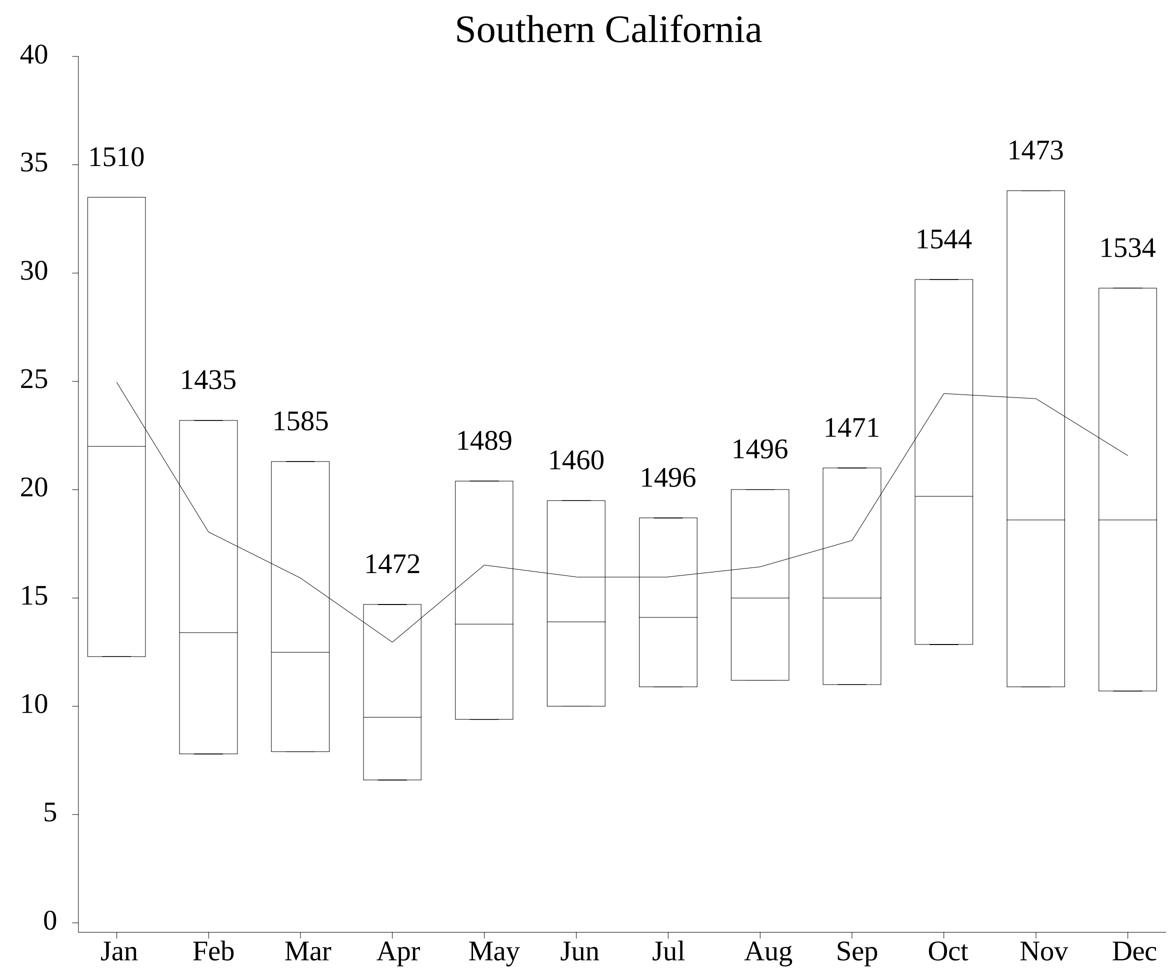
Urban 24-hour average $PM_{2.5}$ concentration distributions by region and month, 2001-2003.



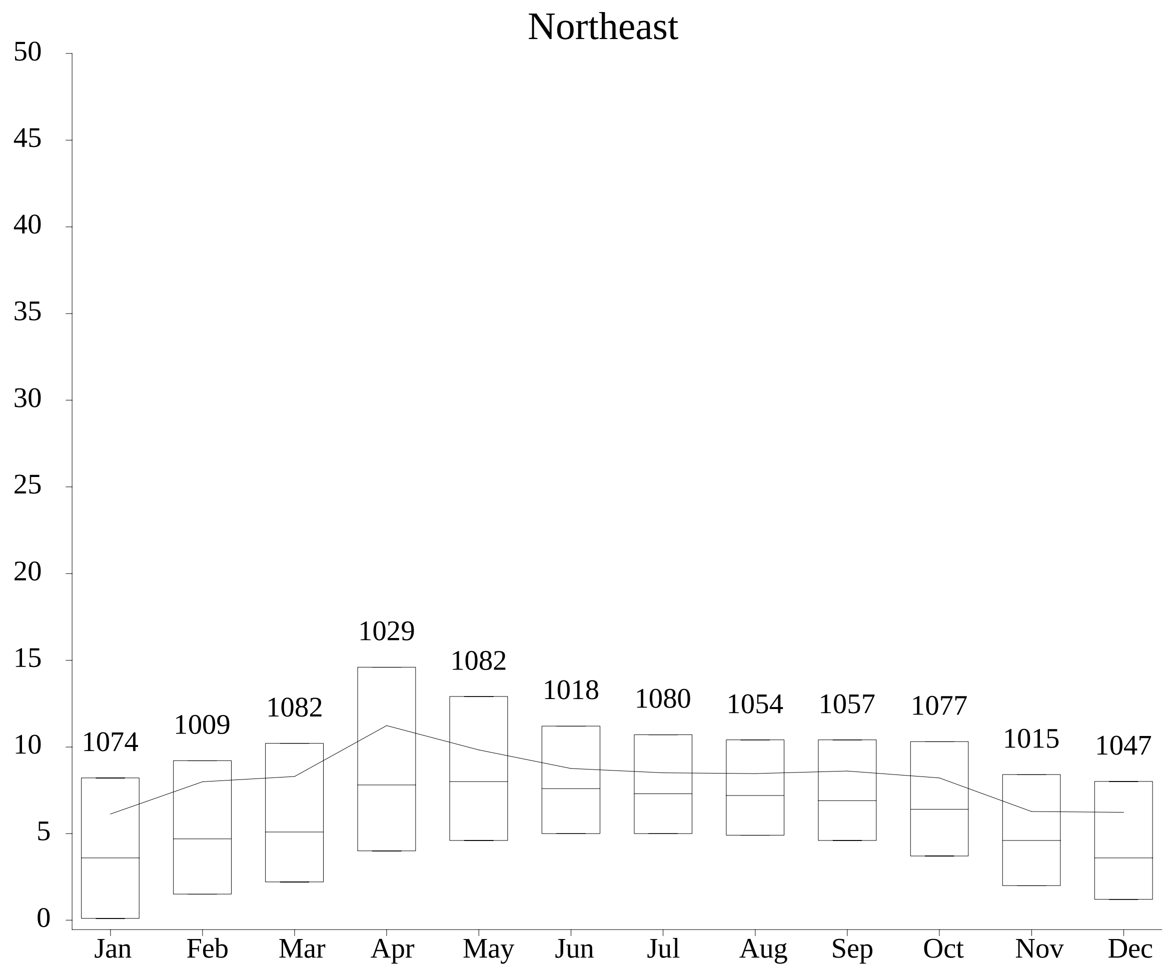
Urban 24-hour average $PM_{2.5}$ concentration distributions by region and month, 2001-2003.



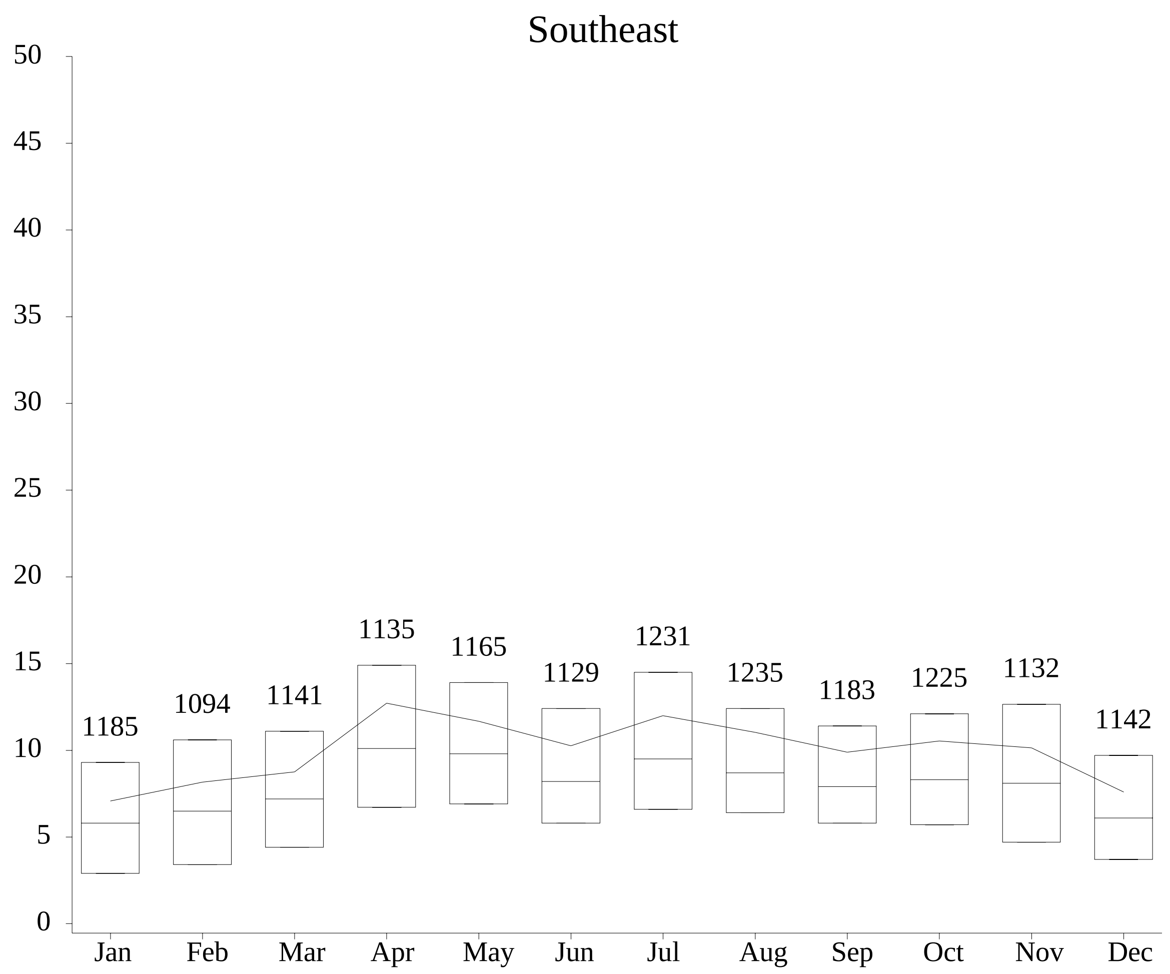
Urban 24-hour average $PM_{2.5}$ concentration distributions by region and month, 2001-2003.



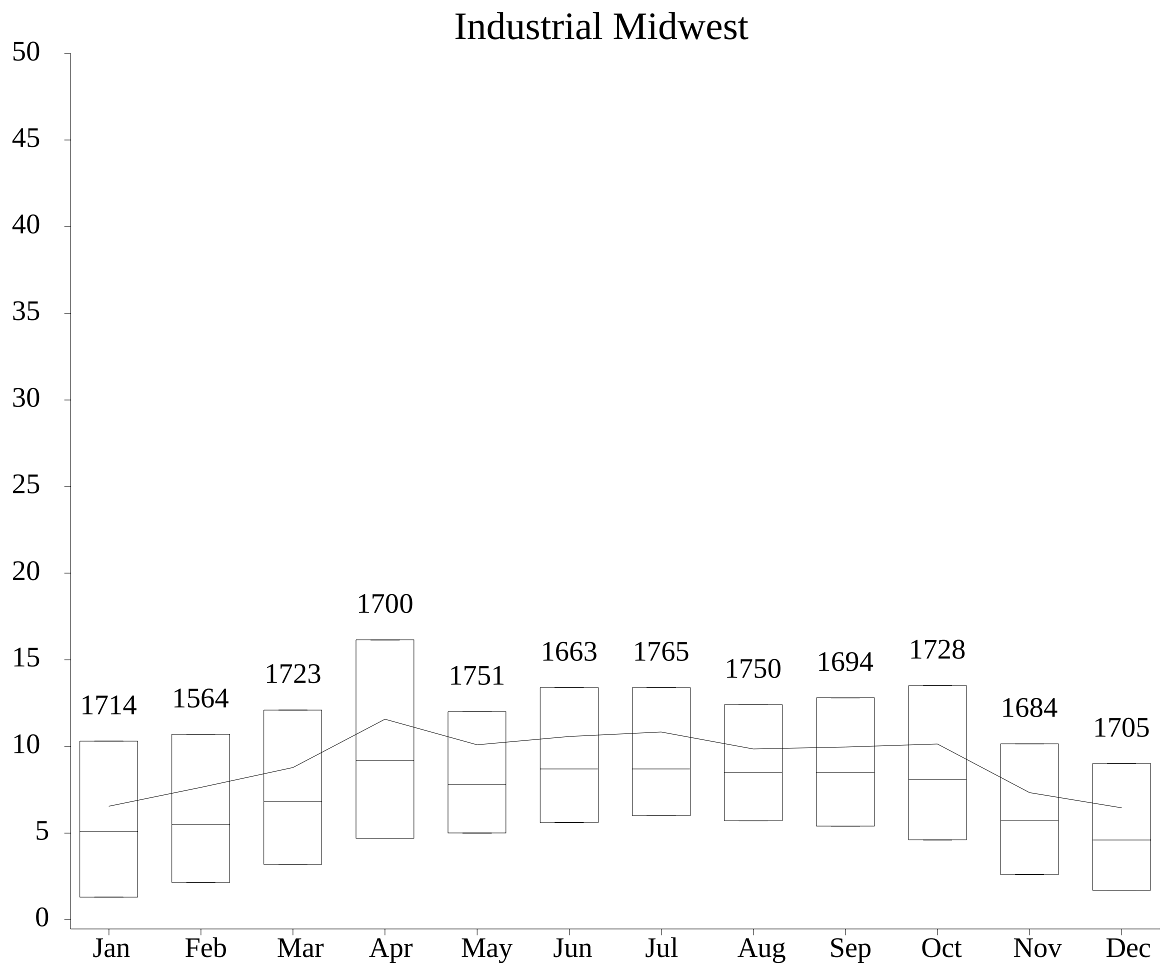
Urban 24-hour average $PM_{10-2.5}$ concentration distributions by region and month, 2001-2003.



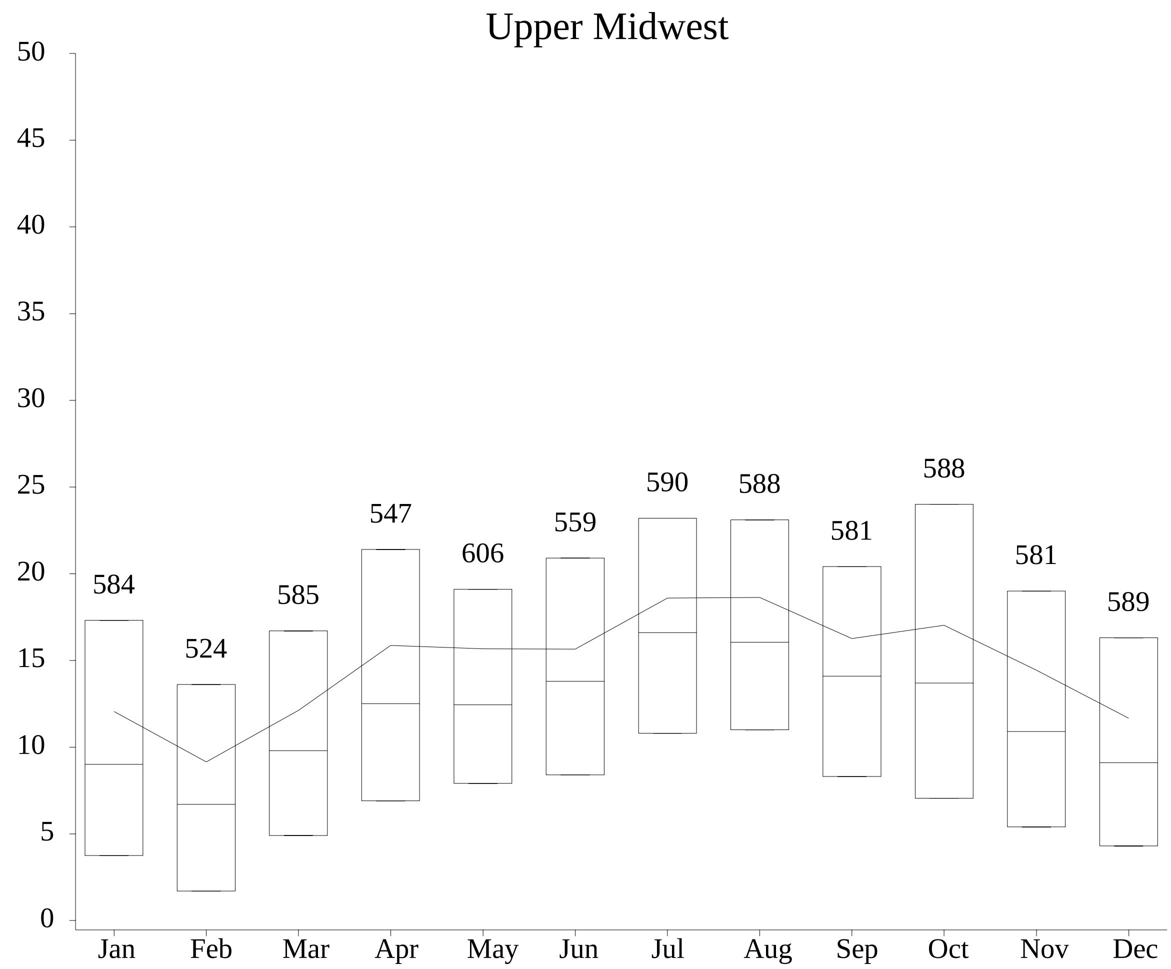
Urban 24-hour average $PM_{10\text{-}2.5}$ concentration distributions by region and month, 2001-2003.



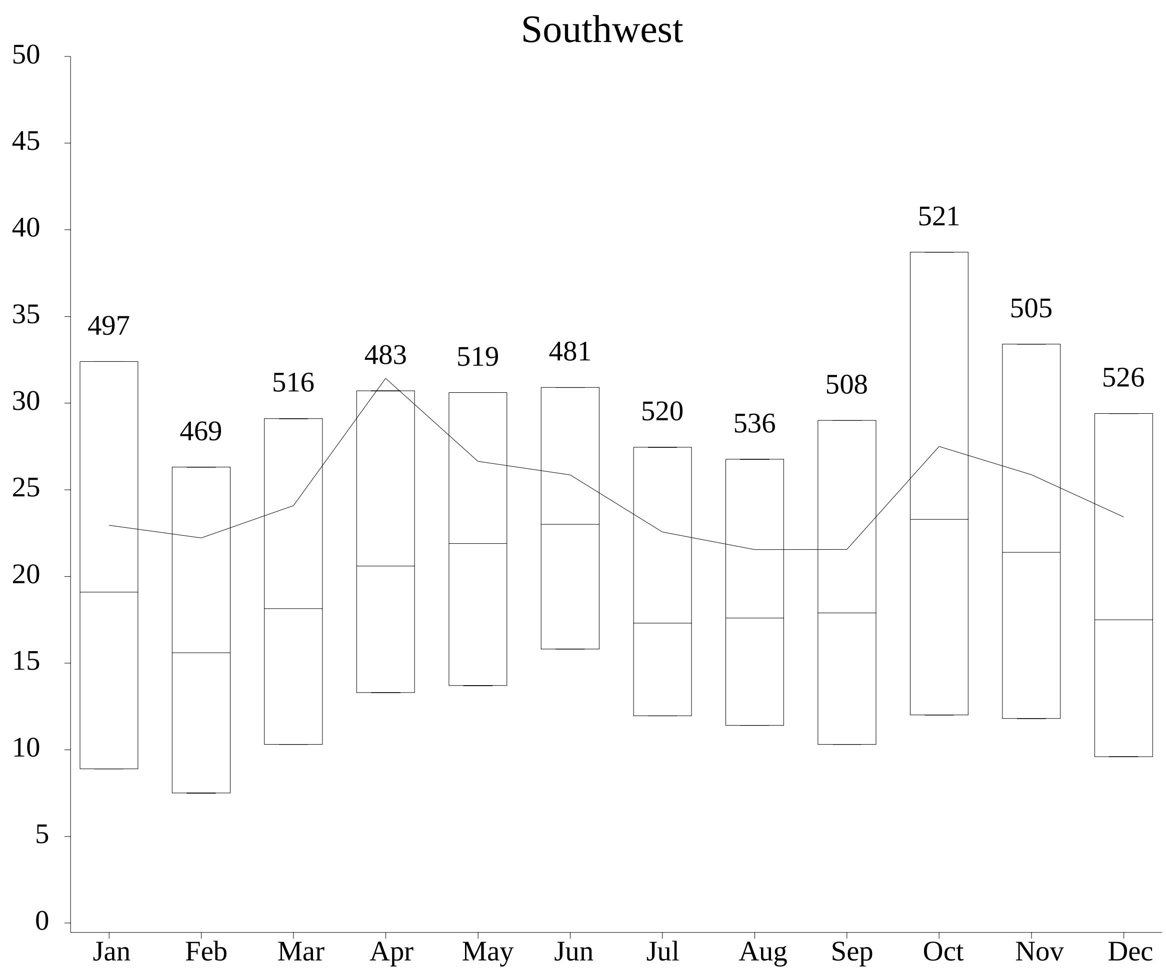
Urban 24-hour average $PM_{10\text{-}2.5}$ concentration distributions by region and month, 2001-2003.



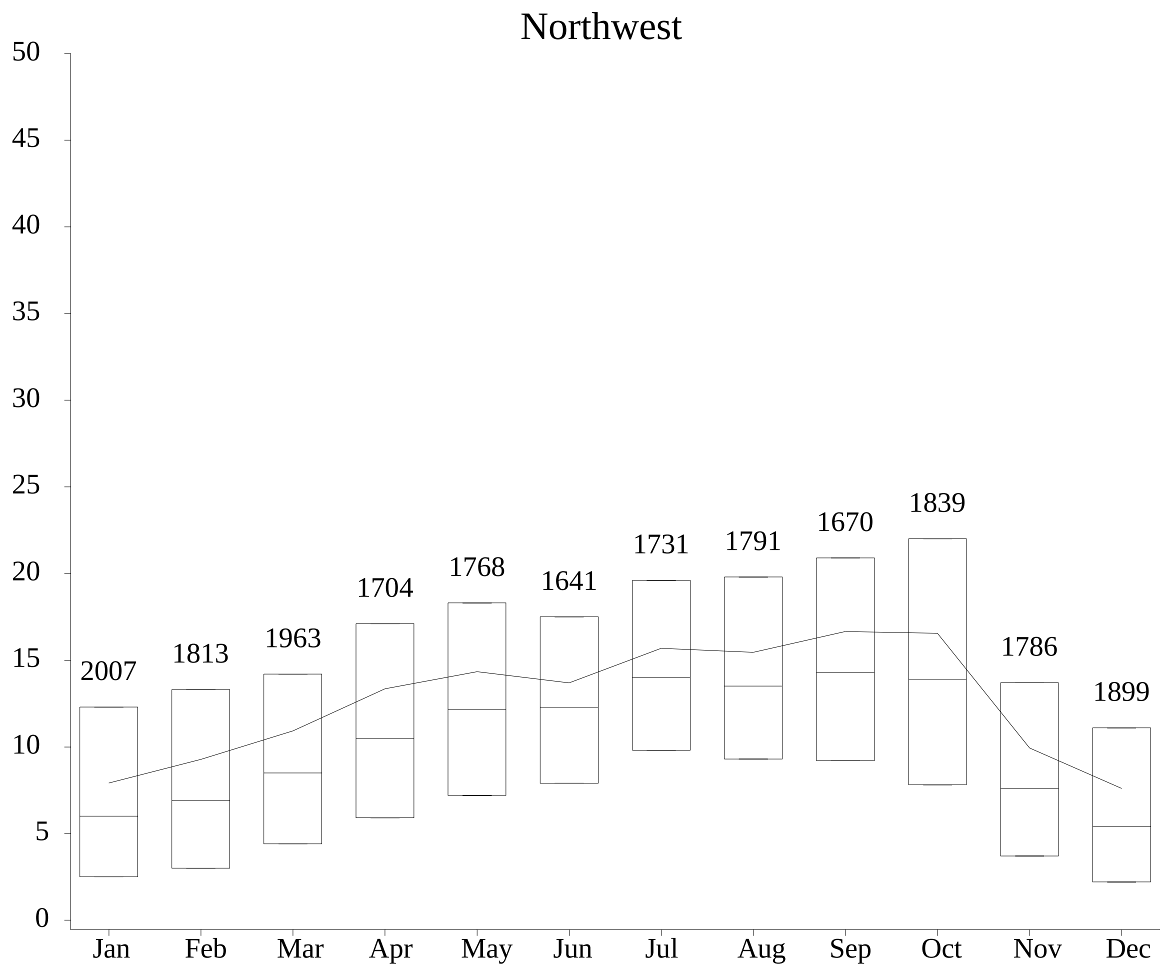
Urban 24-hour average $PM_{10-2.5}$ concentration distributions by region and month, 2001-2003.



Urban 24-hour average $PM_{10-2.5}$ concentration distributions by region and month, 2001-2003.



Urban 24-hour average $PM_{10\text{-}2.5}$ concentration distributions by region and month, 2001-2003.



Urban 24-hour average $PM_{10\text{-}2.5}$ concentration distributions by region and month, 2001-2003.