UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NC 27711

JUL 2 2 2010

OFFICE OF
AIR QUALITY PLANNING
AND STANDARDS

MEMORANDUM

SUBJECT: $PM_{2.5}$ Air Quality Analyses

FROM: Mark Schmidt, OAR/OAQPS/AQAD
Beth Hassett-Sipple, OAR/OAQPS/HEID
Pradeep Rajan, OAR/OAQPS/HEID

TO: PM NAAQS Review (Docket EPA-OAQR-2007-0492)

Overview

This memorandum documents $PM_{2.5}$ air quality analyses conducted for the review of the particulate matter (PM) national ambient air quality standards (NAAQS). The purpose of these analyses was to inform staff conclusions on alternative $PM_{2.5}$ standards that are appropriate to consider in the current PM NAAQS review, specifically focusing on two basic elements of the standards: form[1] and level. This work is in addition to air quality analyses previously conducted and discussed in Hassett-Sipple et al. (2010).[2]

Specific analyses conducted, including tasks, assumptions, caveats, and processing methodologies, are described in more details below. In summary, these analyses address the following:

Annual Standard Analyses

- Analysis 1 – Evaluation of the spatial averaging provisions for the $PM_{2.5}$ annual standard, specifically, a demographic analysis of the potential for disproportionate impacts on potentially vulnerable populations (i.e., environmental justice analysis)
- Analysis 2 – Distributions of air quality and associated population data from selected long- and short-term epidemiological studies

[1] The "form" of a standard defines the air quality statistic that is to be compared to the level of the standard in determining whether an area attains the standard.

[2] Hassett-Sipple, B., Schmidt, M., and Rajan, P. (2010). Analyses of $PM_{2.5}$ Data for the PM NAAQS Review. Memorandum to PM NAAQS review docket. Docket ID number EPA-HQ-OAR-2007-0492-0077. March 29, 2010. Available: http://www.epa.gov/ttn/naaqs/standards/pm/data/20100330dataanalysis.pdf.

24-hour Standard Analyses

- Analysis 3 – Evaluation of "high" $PM_{2.5}$ concentrations - comparison of 98th versus 99th percentiles forms for the 24-hour standard
- Analysis 4 – Distributions of the ratios of 24-hour standard design values (DVs) to annual standard DVs, by geographic region

Suite of Standards Analysis

- Analysis 5 – Assessment of potential $PM_{2.5}$ nonattainment areas[3]

General Data Processing

The air quality data for this project originated from EPA's Air Quality System (AQS) database, the official repository of NAAQS-comparable ambient measurements. Two sets of AQS data were extracted on May 9, 2010: 1) a file of raw 2000-2008, 24-hour duration (filter-based), monitor-level, federal reference method (FRM) data, and 2) a file of hourly (continuous) 2000-2008, monitor-level, federal equivalent method (FEM) data summarized in AQS to a 24-hour basis. The two monitor-level files were subsequently combined and then aggregated to a site basis by averaging by site-day. That is, in situations where there was more than one 24-hour average concentration reported for the same site location (i.e., collocated monitors) for the same day, the multiple 24-hour averages were averaged together. The following statistical metrics were computed from the site-level 24-hour average $PM_{2.5}$ concentrations:

1) annual 98th percentile 24-hour average concentrations,
2) 3-year average 98th percentile concentrations (24-hour design value estimates),
3) annual means, and
4) 3-year average annual means (annual design value estimates).

The metrics were computed at the site-level and then aggregated using the regular (non-seasonal) protocols specified in 40CFR Part 50, Appendix N. A completeness criterion of a minimum of 11 samples per quarter for all 4 quarters of each year (or all 12 quarters of a 3-year period for design value estimates) was imposed on all the site-based metrics. Appendix N provides additional options for validating the annual and 3-year metrics, depending on the metric type and level in relation to the associated NAAQS. To avoid the level and metric validation bias, an across-the-board 11 sample minimum per quarter criterion was imposed. Also, Appendix N has a special 98th percentile computation protocol for sites authorized to sample seasonally (that is, to sample at a less stringent frequency in months when lower concentrations are expected). The seasonal calculation protocol was not used in these analyses; 98th percentile values were computed at all sites using the regular calculation protocol. Because of the possible differences in metric validation and 98th percentile calculation protocol options, the generated 3-year metrics are referred to as design value *estimates*.

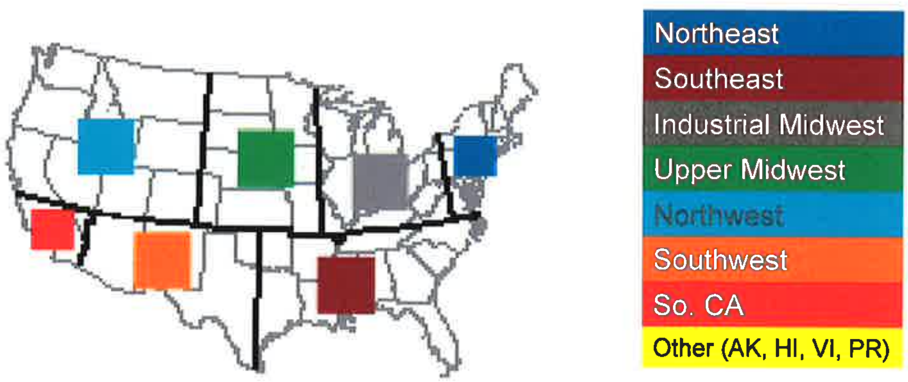
In most of the air quality analyses described below, sites and/or counties were identified and/or aggregated by U.S. geographic region (Figure 1). These regional definitions are identical

[3] This assessment was not considered as a basis for staff conclusions presented in the *Policy Assessment for the Review of the Particulate Matter National Ambient Air Quality Standards: Second External Review Draft* (June 2010), available: http://www.epa.gov/ttn/naaqs/standards/pm/s_pm_2007_pa.html

to regions considered in the *Quantitative Health Risk Assessment for Particulate Matter* (EPA-452/R-10-005)[4] and in previous PM NAAQS reviews.

Additional data processing details are described in the analysis-specific descriptions below.

Figure 1. Regional definitions used in analyses



<u>Analysis 1 – Evaluation of the spatial averaging provisions for the $PM_{2.5}$ annual standard, specifically, a demographic analysis of the potential for disproportionate impacts on potentially vulnerable populations (i.e., environmental justice analysis)</u>

This analysis focused on determining if the spatial averaging provisions as modified in 2006 could introduce inequities in protection for susceptible populations exposed to $PM_{2.5}$. The current form of the annual $PM_{2.5}$ standard includes provisions for spatial averaging if certain criteria are met (i.e., 40 CFR Pt. 50 App. N, 2.0(b)). This analysis looked beyond areas that would meet the spatial averaging criteria and considered all urban areas with at least two valid annual DV monitors. We evaluated whether persons with a lower socioeconomic status (SES) are more likely than the general population to live in areas in which the monitors recording the highest air quality values in an area are located. Data used in this analysis included demographic parameters measured at the Census Block or Census Block Group level, including percent minority population, percent minority subgroup population, percent people living below poverty level, percent people 18 years old or younger, and percent people 65 years and older.

Air quality data from 2006-2008, summarized as 3-year annual $PM_{2.5}$ DV estimates, were evaluated in order to identify the highest concentration site location for each urban area.. Using 2000 Census Block and Block Group information, the population demographics of the areas surrounding the high site location as characterized by buffer radii of 0.5, 1.0, 2.0, and 3.0 miles, were compared to the population demographics for the overall urban area to determine if there were any SES-related differences. Core Based Statistical Area (CBSA) definitions were used to classify the urban areas. As described in the General Data Processing section above, an across-the-board data completeness criterion of 11 or more samples for each quarter of the 12-quarter period (2006-2008) was used to validate the annual DV estimates. One hundred sixteen (116)

[4] Available: http://www.epa.gov/ttn/naaqs/standards/pm/data/PM_RA_FINAL_June_2010.pdf

CBSAs, with sum population of almost 176 million, contained at least two valid annual DV locations for the period 2006-2008 and, hence, were used in the analysis. Table 1 lists the 116 CBSAs and Figure 6 maps them along with the corresponding high site locations. The high site buffer radii demographic attributes were determined by summarizing data for Census Blocks whose centroids fell within the determined cut point distances. The overall CBSA attributes were based on a summarization of all Census Blocks located in the county-based CBSA definitions. Some of the summarized Census population information (i.e., population below the poverty line, population under age 18, and population age 65 and over) were actually reported using Census Block Group as the lowest level, and for these variables, data were prorated to blocks according to total population.

Table 2 shows summary results of the comparison between the four high site buffers and the overall CBSA area. In general, the areas surrounding the high monitor of an urban area were observed to have higher percentages of minorities and people under the poverty level than the urban area as a whole.

Table 1. CBSAs Used in the Spatial Averaging Analysis

CBSA	CBSA name	CBSA population (1000's)	Number of sites in CBSA
10420	Akron, OH	695	3
10740	Albuquerque, NM	730	3
10900	Allentown-Bethlehem-Easton, PA-NJ	740	2
11260	Anchorage, AK	320	2
12060	Atlanta-Sandy Springs-Marietta, GA	4,248	8
12260	Augusta-Richmond County, GA-SC	500	3
12540	Bakersfield, CA	662	4
12580	Baltimore-Towson, MD	2,553	7
12940	Baton Rouge, LA	706	5
13820	Birmingham-Hoover, AL	1,052	10
14460	Boston-Cambridge-Quincy, MA-NH	4,391	10
14500	Boulder, CO	291	2
14860	Bridgeport-Stamford-Norwalk, CT	883	4
15380	Buffalo-Niagara Falls, NY	1,170	3
16580	Champaign-Urbana, IL	210	2
16620	Charleston, WV	310	2
16700	Charleston-North Charleston, SC	549	2
16740	Charlotte-Gastonia-Concord, NC-SC	1,330	4
16860	Chattanooga, TN-GA	477	3
16980	Chicago-Naperville-Joliet, IL-IN-WI	9,098	27
17140	Cincinnati-Middletown, OH-KY-IN	2,010	11
17300	Clarksville, TN-KY	232	2
17460	Cleveland-Elyria-Mentor, OH	2,148	9
17900	Columbia, SC	647	3
17980	Columbus, GA-AL	282	4
18140	Columbus, OH	1,613	2
19100	Dallas-Fort Worth-Arlington, TX	5,162	6
19340	Davenport-Moline-Rock Island, IA-IL	376	4
19380	Dayton, OH	848	3
19740	Denver-Aurora, CO	2,203	5
19780	Des Moines-West Des Moines, IA	481	2
19820	Detroit-Warren-Livonia, MI	4,453	11
20100	Dover, DE	127	2
20260	Duluth, MN-WI	275	3
20940	El Centro, CA	142	3
21340	El Paso, TX	680	3
21660	Eugene-Springfield, OR	323	3
21780	Evansville, IN-KY	343	4
22900	Fort Smith, AR-OK	273	2
23420	Fresno, CA	799	2
23540	Gainesville, FL	232	2
24540	Greeley, CO	181	2
24860	Greenville-Mauldin-Easley, SC	560	2
25180	Hagerstown-Martinsburg, MD-WV	223	2
25420	Harrisburg-Carlisle, PA	509	2
26180	Honolulu, HI	876	3
26420	Houston-Sugar Land-Baytown, TX	4,715	3
26580	Huntington-Ashland, WV-KY-OH	289	2
26900	Indianapolis-Carmel, IN	1,525	5
27260	Jacksonville, FL	1,123	2
27540	Jasper, IN	53	3
28060	Kalispell, MT	74	2
28140	Kansas City, MO-KS	1,836	8
28700	Kingsport-Bristol-Bristol, TN-VA	298	2
28940	Knoxville, TN	616	5
29180	Lafayette, LA	239	2
29340	Lake Charles, LA	194	2
29740	Las Cruces, NM	175	2

CBSA name	CBSA population (1000's)	Number of sites in CBSA
Las Vegas-Paradise, NV	1,376	3
Lexington-Fayette, KY	408	2
Little Rock-North Little Rock-Conway,	611	3
Logan, UT-ID	103	2
Los Angeles-Long Beach-Santa Ana, CA	12,366	10
Louisville/Jefferson County, KY-IN	1,162	7
Macon, GA	222	2
Medford, OR	181	2
Memphis, TN-MS-AR	1,205	4
Miami-Fort Lauderdale-Pompano Beac	5,008	8
Milwaukee-Waukesha-West Allis, WI	1,501	7
Minneapolis-St. Paul-Bloomington, MN	2,969	9
Mobile, AL	400	2
Nashville-Davidson-Murfreesboro-Fran	1,312	3
New Haven-Milford, CT	824	5
New Orleans-Metairie-Kenner, LA	1,317	2
New York-Northern New Jersey-Long	18,323	21
Ogden-Clearfield, UT	443	3
Oklahoma City, OK	1,095	2
Omaha-Council Bluffs, NE-IA	767	5
Orlando-Kissimmee, FL	1,645	3
Oxnard-Thousand Oaks-Ventura, CA	753	4
Philadelphia-Camden-Wilmington, PA-	5,687	12
Phoenix-Mesa-Scottsdale, AZ	3,252	8
Pittsburgh, PA	2,431	11
Portland-Vancouver-Beaverton, OR-W	1,928	2
Providence-New Bedford-Fall River, RI	1,583	5
Provo-Orem, UT	377	4
Rapid City, SD	113	2
Richmond, VA	1,097	4
Riverside-San Bernardino-Ontario, CA	3,255	9
Sacramento–Arden-Arcade–Roseville,	1,797	5
St. Louis, MO-IL	2,721	9
Salt Lake City, UT	969	4
San Diego-Carlsbad-San Marcos, CA	2,814	4
San Francisco-Oakland-Fremont, CA	4,124	5
San Juan-Caguas-Guaynabo, PR	2,509	2
San Luis Obispo-Paso Robles, CA	247	2
Santa Barbara-Santa Maria-Goleta, CA	399	2
Savannah, GA	293	2
Seattle-Tacoma-Bellevue, WA	3,044	2
Sheridan, WY	27	2
Springfield, MA	680	3
Tampa-St. Petersburg-Clearwater, FL	2,396	4
Terre Haute, IN	171	2
Toledo, OH	659	2
Trenton-Ewing, NJ	351	2
Truckee-Grass Valley, CA	92	2
Tucson, AZ	844	2
Tulsa, OK	860	2
Virginia Beach-Norfolk-Newport News	1,576	3
Washington-Arlington-Alexandria, DC-	4,796	9
Weirton-Steubenville, WV-OH	132	5
Wheeling, WV-OH	153	2
Wichita, KS	571	4
Winston-Salem, NC	422	2
Worcester, MA	751	2
Youngstown-Warren-Boardman, OH-P	603	4
Total population of 116 CBSAs	**175,761**	

Figure 2. Map of CBSAs Used in Spatial Averaging Analysis.

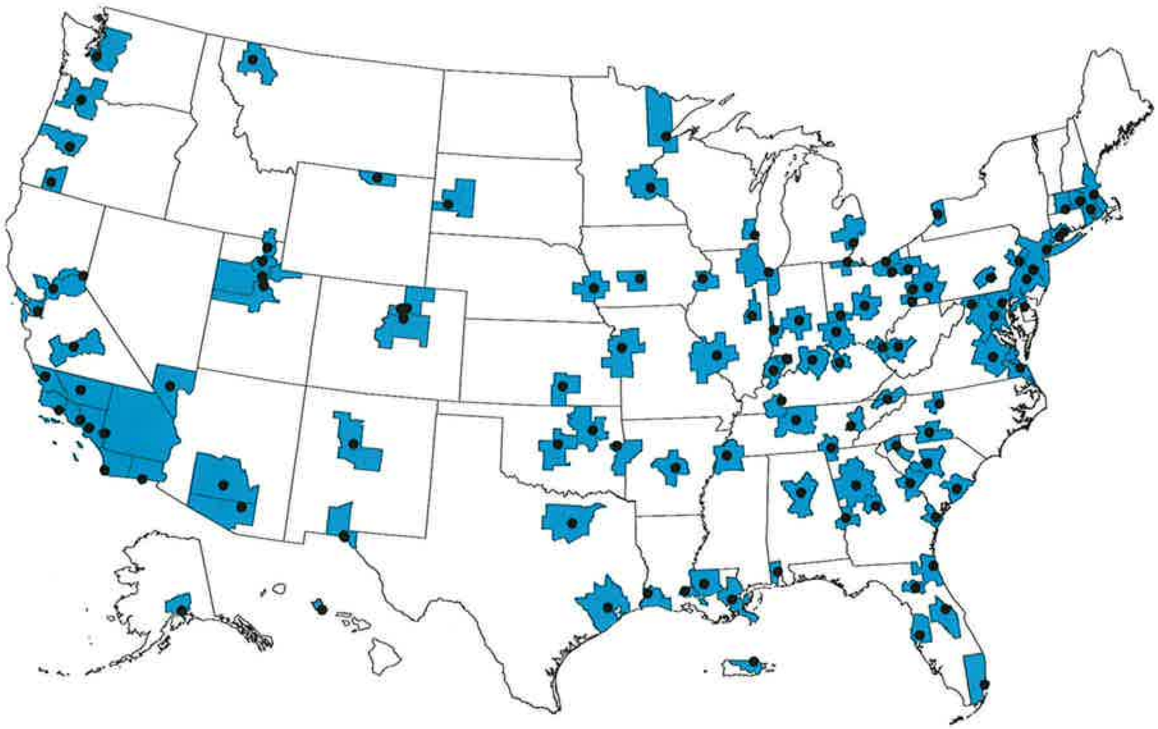


Table 2. Summary Results: Comparison Between the Four High Site Buffers and the Overall CBSA Area

Area counts (for 116 CBSAs) comparing SES variable within 0.5, 1.0, 2.0, and 3.0 mile radii buffers versus within overall CBSA

Population Variable	Buffer	Percentage of areas where variable percentage greater in high site buffer than in overall CBSA	Percentage of areas where variable percentage not greater in high site buffer than in overall CBSA	Mean difference in variable percentage in high site buffer minus variable percentage in overall CBSA	Median difference in variable percentage in high site buffer minus variable percentage in overall CBSA	Mean ratio of variable percentage in high site buffer / variable percentage in overall CBSA	Median ratio of variable percentage in high site buffer / variable percentage in overall CBSA
Minority	0.5 mile	58.9%	41.1%	8.9%	2.8%	1.57	1.16
Minority	1.0 mile	67.2%	32.8%	11.6%	6.1%	1.66	1.37
Minority	2.0 mile	72.4%	27.6%	11.6%	7.8%	1.63	1.50
Minority	3.0 mile	76.7%	23.3%	11.4%	9.6%	1.59	1.47
Minority	avg of 4 buffers	68.8%	31.2%	10.9%	6.6%	1.61	1.38
African American	0.5 mile	49.1%	50.9%	6.5%	-0.1%	1.83	0.90
African American	1.0 mile	57.8%	42.2%	7.3%	1.1%	1.66	1.16
African American	2.0 mile	58.6%	41.4%	8.2%	1.0%	1.62	1.32
African American	3.0 mile	66.4%	33.6%	8.4%	3.1%	1.59	1.52
African American	avg of 4 buffers	58.0%	42.0%	7.6%	1.3%	1.68	1.23
Native American	0.5 mile	52.7%	47.3%	0.3%	0.0%	1.51	1.03
Native American	1.0 mile	62.9%	37.1%	0.8%	0.1%	1.64	1.21
Native American	2.0 mile	65.5%	34.5%	0.5%	0.1%	1.43	1.14
Native American	3.0 mile	66.4%	33.6%	0.5%	0.0%	1.33	1.11
Native American	avg of 4 buffers	61.9%	38.1%	0.5%	0.1%	1.48	1.12
Other and Multiracial	0.5 mile	59.8%	40.2%	2.1%	0.4%	1.32	1.10
Other and Multiracial	1.0 mile	58.6%	41.4%	3.5%	2.1%	1.46	1.25
Other and Multiracial	2.0 mile	66.4%	33.6%	2.9%	1.0%	1.37	1.19
Other and Multiracial	3.0 mile	66.4%	33.6%	2.5%	1.2%	1.32	1.27
Other and Multiracial	avg of 4 buffers	62.8%	37.2%	2.8%	1.2%	1.37	1.20
Hispanic	0.5 mile	55.4%	44.6%	3.9%	0.3%	1.50	1.09
Hispanic	1.0 mile	60.3%	39.7%	5.8%	1.1%	1.77	1.28
Hispanic	2.0 mile	66.4%	33.6%	5.3%	1.4%	1.68	1.32
Hispanic	3.0 mile	75.0%	25.0%	4.5%	1.1%	1.54	1.33
Hispanic	avg of 4 buffers	64.3%	35.7%	4.9%	1.0%	1.62	1.26
Below Poverty Line	0.5 mile	75.0%	25.0%	0.1%	0.0%	1.74	1.39
Below Poverty Line	1.0 mile	75.9%	24.1%	0.1%	0.1%	1.77	1.51
Below Poverty Line	2.0 mile	77.6%	22.4%	0.1%	0.1%	1.70	1.60
Below Poverty Line	3.0 mile	76.7%	23.3%	0.1%	0.1%	1.59	1.55
Below Poverty Line	avg of 4 buffers	76.3%	23.7%	0.1%	0.1%	1.70	1.51
Under age 18	0.5 mile	48.2%	51.8%	-1.6%	-0.5%	0.94	0.98
Under age 18	1.0 mile	50.9%	49.1%	-0.2%	0.1%	0.99	1.01
Under age 18	2.0 mile	51.7%	48.3%	0.0%	0.0%	1.00	1.02
Under age 18	3.0 mile	54.3%	45.7%	0.1%	0.2%	1.00	1.01
Under age 18	avg of 4 buffers	51.3%	48.7%	-0.4%	-0.1%	0.98	1.01
Age 65 and over	0.5 mile	53.6%	46.4%	-0.5%	0.3%	0.97	1.03
Age 65 and over	1.0 mile	49.1%	50.9%	0.0%	-0.2%	1.01	0.98
Age 65 and over	2.0 mile	52.6%	47.4%	0.2%	0.1%	1.00	1.01
Age 65 and over	3.0 mile	56.0%	44.0%	0.3%	0.3%	1.04	1.03
Age 65 and over	avg of 4 buffers	52.8%	47.2%	0.0%	0.1%	1.01	1.01

Average percentage of people living within each buffer for each demographic category

Area / buffer	Minority	Native American	Other and Multiracial	Hispanic	Living Below Poverty Line	Under Age 18	Age 65 and Over
CBSA	22.9	0.8	9.9	11.1	0.1	27.3	12.0
within half mile	32.9	1.2	12.4	15.5	0.2	26.6	12.0
within 1 mile	34.4	1.6	13.3	16.9	0.2	27.1	12.0
within 2 miles	34.5	1.3	12.8	16.4	0.2	27.3	12.2
within 3 miles	34.2	1.3	12.4	15.5	0.2	27.3	12.4

Analysis 2 – Distributions of air quality and associated population data from selected long- and short-term epidemiological studies

This analysis built on previous air quality analyses described in an earlier EPA staff memorandum (Hassett-Sipple, et al., 2010).[5] To maintain continuity and consistency, the same air quality database was utilized for all of analyses. Specifically, the EPA database for this work emanated from the site-based $PM_{2.5}$ FRM summary files processed for official DV calculations (circa. July 7, 2009) based on raw data reported in the AQS.[6,7]

Hassett-Sipple et al. (2010), reported NAAQS-like air quality metrics for the same (or similar) geographic areas and time periods utilized by the study authors of selected epidemiological studies, henceforth, these statistics are referred to as area-period air quality metrics. The EPA computed the area-period air quality metrics (i.e., 98th percentile values and annual means) using two approaches for assessing the air quality distribution: (1) a composite monitor or spatial averaging approach and (2) a maximum value or highest monitor approach. The former method is similar to approaches used by many health researchers. The latter method is more comparable to the metrics used to inform NAAQS attainment decisions. By having both sets of statistics available for the same sets of study areas, we could better understand the air quality distributions considered in the epidemiological studies to place those studies in a policy-relevant context.

In the extended analysis documented here, we considered distributions from epidemiological studies that provide a continuous measure of the density of the data from multi-city studies as a function of the annual mean concentrations across the cities in each study. Specifically, we considered the distributions of air quality data as well as distributions of population data across air quality concentrations, which serve as a surrogate for the density of the health effects data. Populations were estimated for each study by summing the county populations (per 2000 Census) of the counties where sites had valid contributing data for the area-period air quality metric estimates. In Figures 3 through 6, the cumulative variable (i.e., frequency or population) is plotted on the y-axis as a percentage of the total study variable (i.e., as a percentage of the total study area-period count or the total study area-period population). The x-axis represents the annual mean concentration in units of $\mu g/m^3$. In all four figures, horizontal (i.e., y-axis) reference lines are drawn at the 10th, 25th, 50th, 75th, and 90th percentiles; Table 3 identifies the total study counts (number of areas), total study populations, and approximate percentile reference line air quality values for the four figures.

---

[5] Hassett-Sipple, B, Schmidt, M, and Rajan, P. (2010) Analyses of $PM_{2.5}$ Data for the PM NAAQS Review. Memorandum to the PM NAAQS review docket. Docket ID number EPA-HQ-OAR-2007-0492-0077. March 29, 2010. Available: http://www.epa.gov/ttn/naaqs/standards/pm/data/20100330dataanalysis.pdf.

[6] Air quality data and associated processing details are available at: http://epa.gov/airtrends/pdfs/dv_pm25_2006_2008rev102809.xls

[7] The annual air quality metrics did not include consideration of significant, policy-relevant exceptional events, that is, flagged data points that were approved (i.e., "concurred") by an EPA Regional Office. Exceptional events are events for which the normal planning and regulatory process established by the Clean Air Act (CAA) is not appropriate. Section 319 of the CAA defines an event as an exceptional event if the event affects air quality; is an event that is not reasonably controllable or preventable; is an event caused by human activity that is unlikely to recur at a particular location or a natural event; and is determined by EPA to be an exceptional event. The statutory definition of exceptional event specifically excludes stagnation of air masses or meteorological inversions; a meteorological event involving high temperatures or lack of precipitation; or air pollution relating to source noncompliance (72 FR 13561, March 22, 2007).

Figure 3. Air Quality Distributions: Long-term $PM_{2.5}$ Exposure Studies

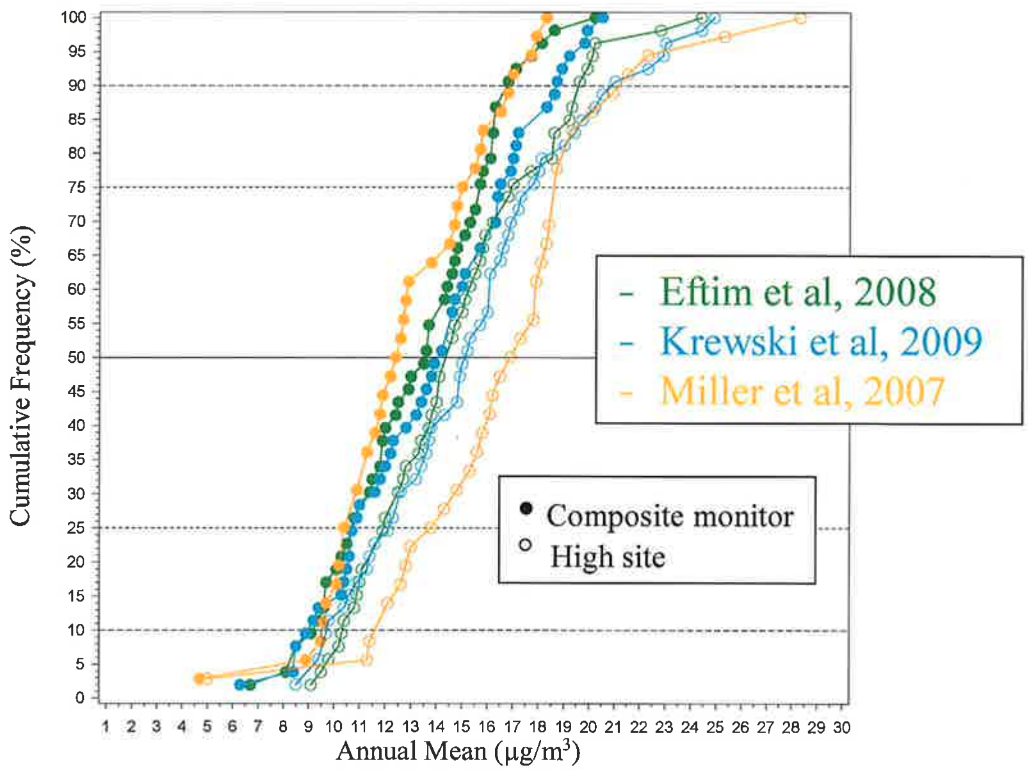


Figure 4. Population Distributions: Long-term $PM_{2.5}$ Exposure Studies

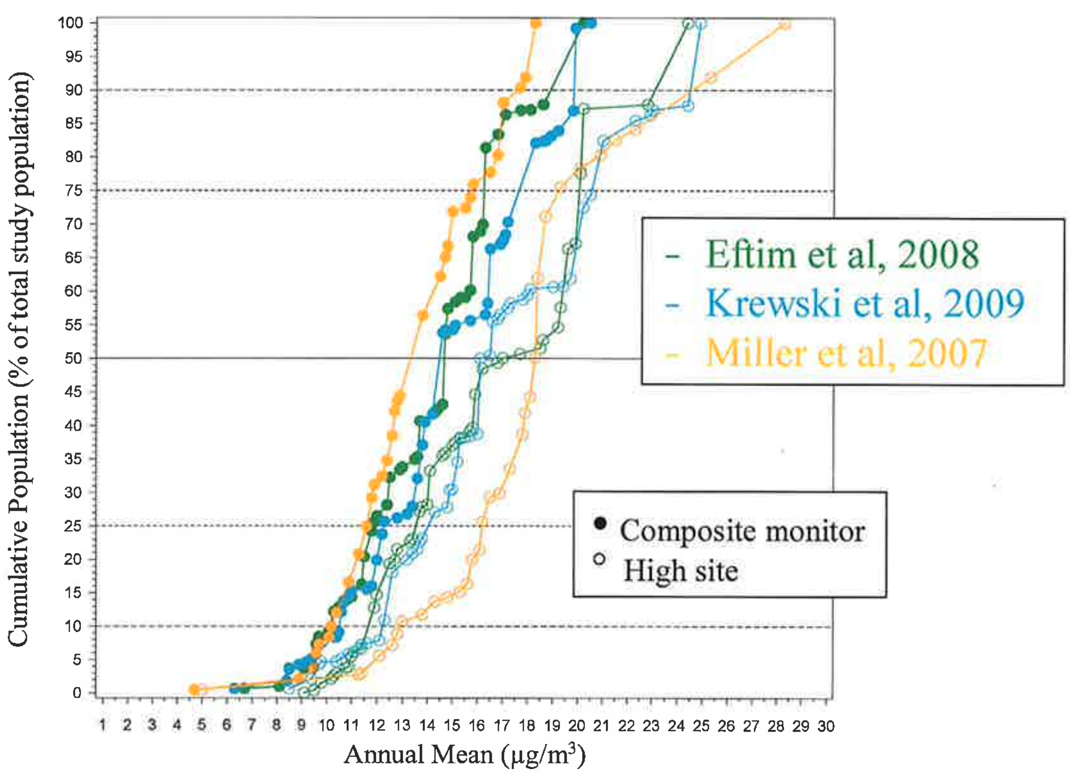


Figure 5. Air Quality Distributions: Short-term $PM_{2.5}$ Exposure Studies

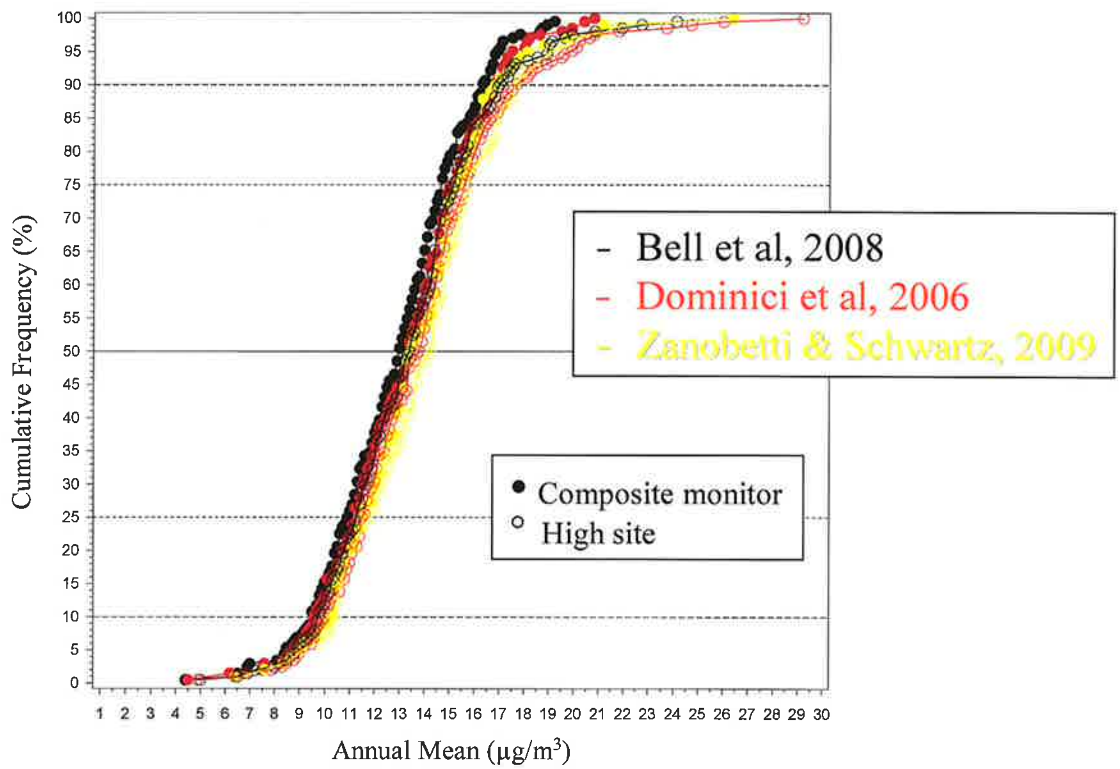


Figure 6. Population Distributions: Short-term $PM_{2.5}$ Exposure Studies

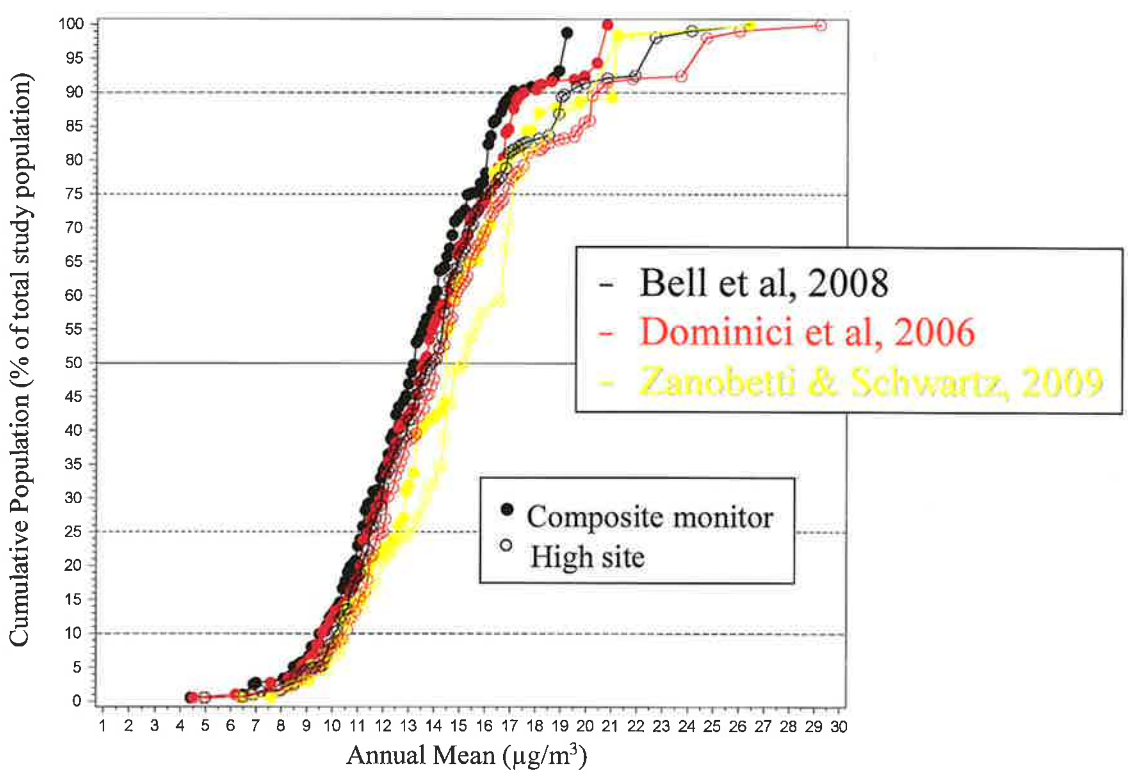


Table 3. Total Study Counts (number of areas), Total Study Populations, and Approximate Percentile Reference Line Air Quality Values for Figure 3, Figure 4, Figure 5, and Figure 6

Figure 3 total study area counts and approximate percentile reference line air quality plotting points

Study / method →	Eftim, composite monitor	Krewski, composite monitor	Miller, composite monitor	Eftim, high site	Krewski, high site	Miller, High site
Total study area counts →	53	53	36	53	53	36
Reference percentile ↓	Approximate air quality value of reference percentile (µg/m³) ↓					
90th	16.7	18.7	16.9	19.5	20.8	21.1
75th	15.6	16.5	15.0	16.9	17.4	18.6
50th	13.5	14.1	12.4	14.3	15.1	16.9
25th	10.7	10.7	10.4	11.9	12.1	13.8
10th	9.1	9.0	9.5	10.3	9.7	11.7

Figure 4 total study area populations and approximate percentile reference line air quality plotting points

Study / method→	Eftim, composite monitor	Krewski, composite monitor	Miller, composite monitor	Eftim, high site	Krewski, high site	Miller, High site
Total study area population (1000s) →	101,439	100,253	248,965	101,439	100,253	248,965
Reference percentile ↓	Approximate air quality value of reference percentile (µg/m³) ↓					
90th	19.2	19.8	17.6	22.0	24.5	24.3
75th	16.2	17.7	15.7	20.0	20.6	19.2
50th	14.7	14.4	13.3	17.0	16.1	18.3
25th	11.9	12.3	11.6	13.5	14.0	16.2
10th	10.2	10.5	10.2	11.7	12.2	12.9

Figure 5 total study area counts and approximate percentile reference line air quality plotting points

Study / method →	Bell, composite monitor	Dominici, composite monitor	Zanobetti and Schwartz, composite monitor	Bell, high site	Dominici, high site	Zanobetti and Schwartz, high site
Total study area counts →	204	204	99	204	204	99
Reference percentile ↓	Approximate air quality value of reference percentile (µg/m³) ↓					
90th	16.3	16.8	16.8	17.0	17.6	17.6
75th	14.7	15.1	15.3	15.1	15.6	15.7
50th	13.0	13.3	13.4	13.4	13.8	14.2
25th	10.9	11.1	11.5	11.2	11.4	11.8
10th	9.5	9.6	10.3	9.8	10.0	10.4

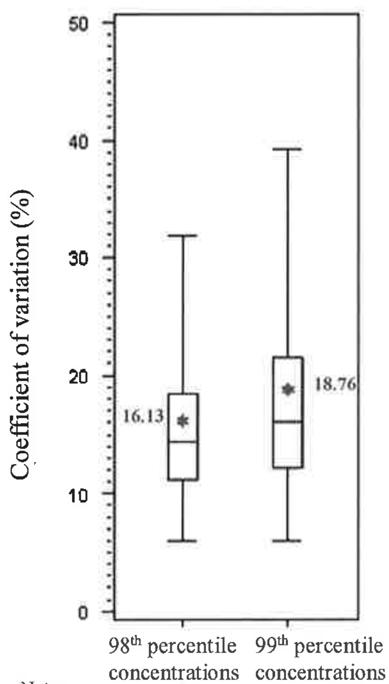
Figure 6 total study area populations and approximate percentile reference line air quality plotting points

Study / method→	Bell, composite monitor	Dominici, composite monitor	Zanobetti and Schwartz, composite monitor	Bell, high site	Dominici, high site	Zanobetti and Schwartz, high site
Total study area population (1000s) →	170,425	168,416	103,285	170,425	168,416	103,285
Reference percentile ↓	Approximate air quality value of reference percentile (µg/m³) ↓					
90th	17.3	17.5	21.0	19.2	20.3	21.0
75th	15.4	16.1	16.3	15.9	16.8	17.1
50th	13.2	13.6	14.5	13.9	14.1	15.2
25th	11.2	11.4	12.5	11.5	12.0	13.1
10th	9.6	9.6	10.3	10.1	10.5	10.6

Analysis 3 – Evaluation of "high" $PM_{2.5}$ concentrations - comparison of 98th versus 99th percentiles forms for the 24-hour standard

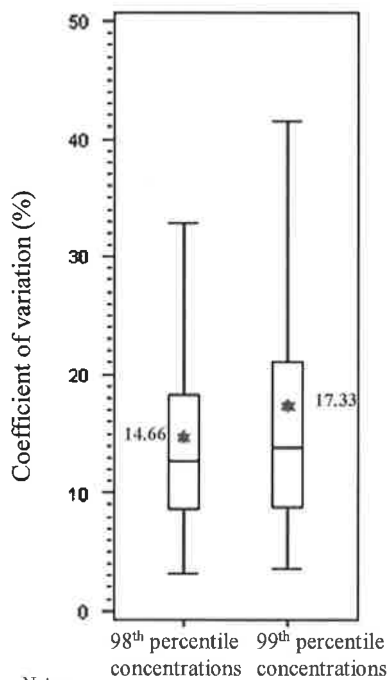
The current form of the 24-hour $PM_{2.5}$ standard is the 98th percentile of 24-hour concentrations at each population-oriented monitor within an area, averaged over 3 years. The relative stability of a 98th percentile form compared to an alternative form based on the 99th percentile values was evaluated. The coefficient of variation was computed for both statistics by site using all valid site-years over a 9-year period (2000-2008), and also for the three most recent years (2006-2008) in which data were available. At least two site-years had to be present in order to derive a valid site-level coefficient of variation. Coefficient of variation was used to normalize the data. Figure 7 shows a box plot of the distribution of site-year coefficient of variation (for site-year 98th and 99th percentile 24-hour concentrations) for the period 2000-2008 and Figure 8 shows the same for the period 2006-2008.

Figure 7. Distribution of site-level variation in 98th and 99th percentile concentrations, as measured by percent coefficient of variation (SD/Mean * 100), computed by site across years, 2000-2008



Notes:
- The box indicates the distribution interquartile range, the line through the box denotes the distribution median, the whisker caps mark the 5th and 95th percentiles, and a blue star identifies the distribution mean.
- Blue values correspond to distribution means

Figure 8. Distribution of site-level variation in 98th and 99th percentile concentrations, as measured by percent coefficient of variation (SD/Mean * 100), computed by site across years, 2006-2008



Notes:
- The box indicates the distribution interquartile range, the line through the box denotes the distribution median, the whisker caps mark the 5th and 95th percentiles, and a blue star identifies the distribution mean.
- Blue values correspond to distribution means

Analysis 4 – Distributions of the ratios of 24-hour DVs to annual DVs, by geographic region

Air quality data from 2000-2008 were evaluated to identify the distributions of the ratios of 98th percentile DVs to annual mean DVs by geographic region. Site-level $PM_{2.5}$ annual mean estimates and annual 98th percentile concentration estimates were calculated using a completeness criterion of 11 or more samples for each of the four quarters per year. The site-year results were aggregated by region. Figures 9 and 10 show box plots of the site-year peak-to-mean ratio distributions for 2000-2008 and 2006-2008, respectively. Figure 11 is a scatter plot of the 2006-2008 site-level 24-hour DV estimates (y-axis) versus the 2006-2008 site-level annual DV estimates (x-axis). As previously noted, only valid DV estimates (i.e., those in which all 12 quarters contained a minimum of 11 samples) were used in the evaluation. Each symbol represents a distinct site and is color-coded according to geographic region. As illustrated in Figure 11, we have examined the implications of a 24-hour standard level of 30 $\mu g/m^3$ as compared to a level of 35 $\mu g/m^3$ (see horizontal reference lines) in conjunction with an alternative standard level of 12 $\mu g/m^3$ (see vertical reference line) to understand where the annual standard would be "generally controlling" compared to areas where the 24-hour standard would be controlling. As illustrated in Figure 11, combining an annual standard of 12 $\mu g/m^3$ with a 24-hour standard of 35 $\mu g/m^3$ results in the 24-hour standard being generally controlling primarily in the Northwest. Alternatively, combining an annual standard of 12 $\mu g/m^3$ with a 24-hour standard of 30 $\mu g/m^3$ would result in many more areas outside of the Northwest in which the 24-hour standard would become the controlling standard. This can be seen by looking to the left of the 12/30 and 12/35 diagonal reference lines, which are the regions in Figure 11 in which the 24-hour standard is estimated to be the controlling standard.

Figure 9. Distribution of Site-level Ratio of Annual 98th Percentile Concentration/Annual Mean, 2000-2008

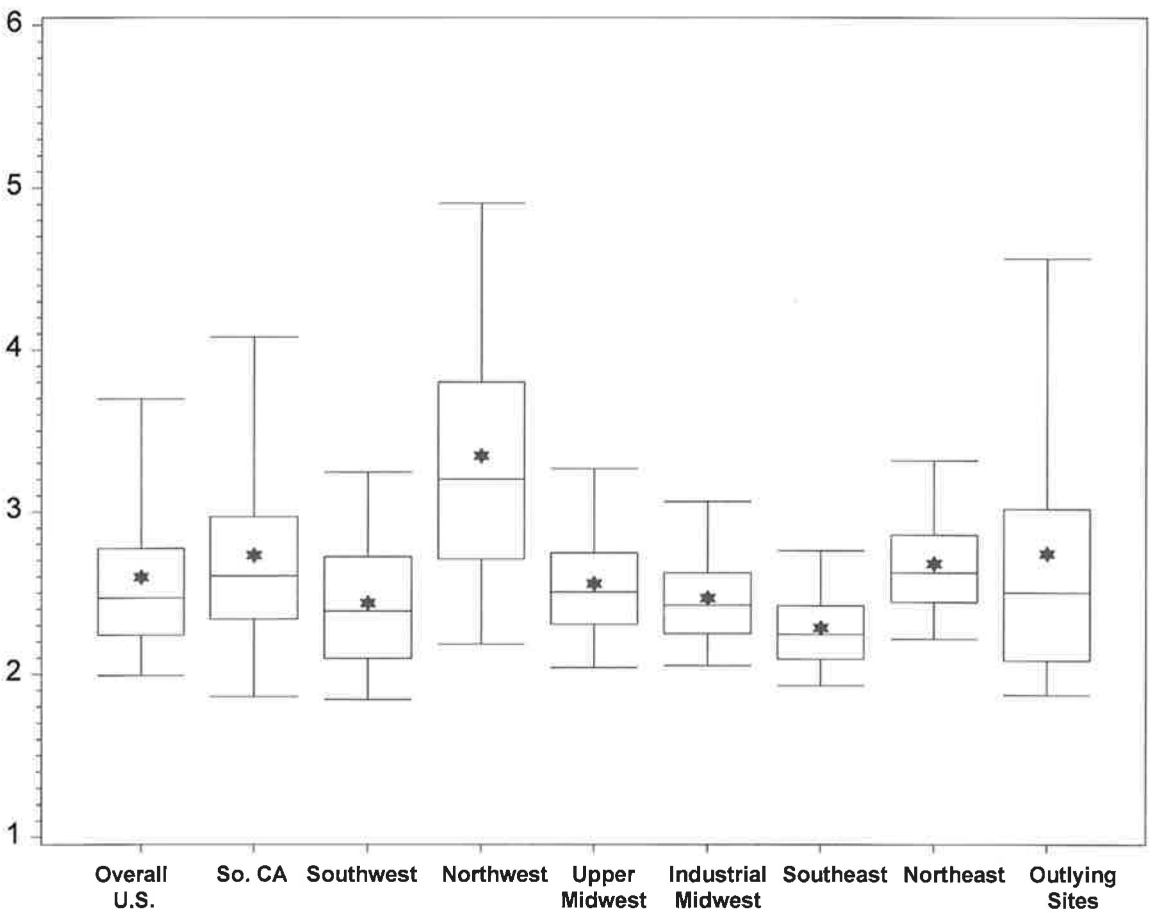


Figure 10. Distribution of Site-level Ratio of Annual 98th Percentile Concentration/Annual Mean, 2006-2008

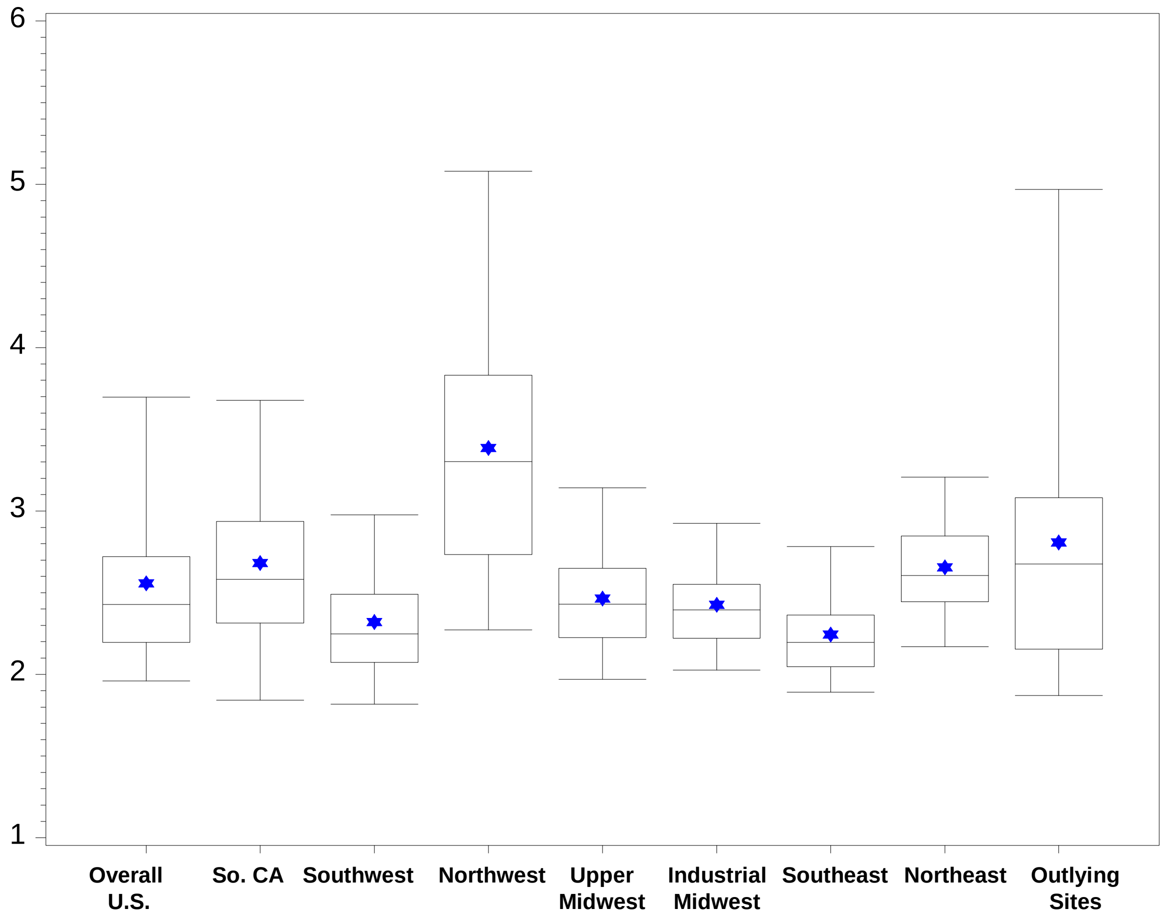
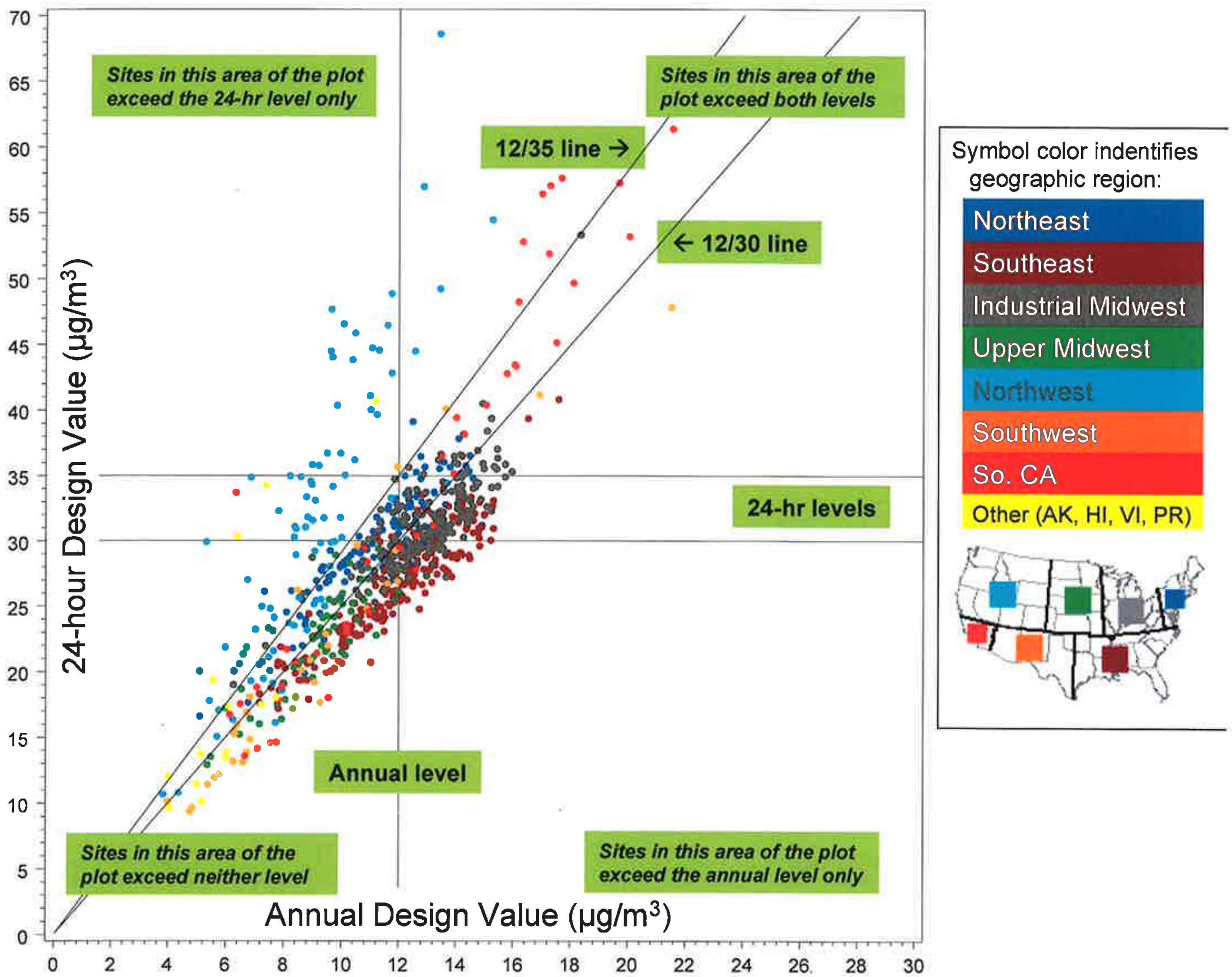


Figure 11. Site-level 24-hour DVs versus Annual DVs, 2006-2008.



<u>Analysis 5 – Assessment of potential $PM_{2.5}$ nonattainment areas</u>

To provide some perspective on the implications of various alternative suites of annual and 24-hour standards, we calculated $PM_{2.5}$ annual and 24-hour DV estimates (i.e., those in which all 12 quarters contained a minimum of 11 samples) at the county-level using the most recent (2006-2008) air quality data available. The percentage of counties and the population in those counties that would likely not attain various alternative suites of standards is presented in Table 4. Results were summarized for the entire U.S. level as well as the specific geographic regions illustrated in Figure 1.

Table 4: Predicted Percent of Counties with Monitors (and Percent of Population in Counties With Monitors) Not Likely to Meet Alternative Annual and 24-hour $PM_{2.5}$ Standard Levels

Region >			All U.S.	Northeast	Southeast	Industrial Midwest	Upper Midwest	Southwest	Northwest	Southern California	Outlying areas
Total # of counties >			526	87	150	134	44	21	63	16	12
Total population (x 1,000)>			182,075	42,841	40,904	38,646	7,361	9,311	18,446	22,610	2,099
Current levels											
annual $\mu g/m^3$	24-hour $\mu g/m^3$	Statistic	Numbers of counties, populations, and percentages of total								
15	35	# counties	62	9	6	15	0	4	19	9	1
		population	47,896	5,639	5,157	8,551	0	1,040	9,847	17,721	83
		% # counties	12%	10%	4%	11%	0%	19%	30%	56%	8%
		% population	26%	13%	13%	22%	0%	11%	53%	78%	4%
Alternative levels											
annual $\mu g/m^3$	24-hour $\mu g/m^3$	Statistic	Numbers of counties, populations, and percentages of total								
13	35	# counties	151	19	45	53	0	4	20	10	1
		population	77,390	10,202	13,678	22,128	0	1,040	9,866	20,535	83
		% # counties	29%	22%	30%	40%	0%	19%	32%	63%	8%
		% population	43%	24%	33%	57%	0%	11%	53%	91%	4%
13	30	# counties	211	40	49	72	0	4	34	11	2
		population	104,539	26,228	14,063	28,417	0	1,040	14,266	20,553	114
		% # counties	40%	46%	33%	54%	0%	19%	54%	69%	17%
		% population	57%	61%	34%	74%	0%	11%	77%	91%	5%
12	35	# counties	239	36	79	90	0	4	20	10	1
		population	102,274	21,725	18,913	30,255	0	1,040	9,866	20,535	83
		% # counties	45%	41%	53%	67%	0%	19%	32%	63%	8%
		% population	56%	51%	46%	78%	0%	11%	53%	91%	4%
12	30	# counties	270	46	79	95	0	4	34	11	2
		population	114,610	27,991	18,913	31,876	0	1,040	14,266	20,553	114
		% # counties	51%	53%	53%	71%	0%	19%	54%	69%	17%
		% population	63%	65%	46%	82%	0%	11%	77%	91%	5%
11	30	# counties	343	53	115	119	5	5	34	11	2
		population	132,785	30,053	26,400	36,179	1,250	4,112	14,266	20,553	114
		% # counties	65%	61%	77%	89%	11%	24%	54%	69%	17%
		% population	73%	70%	65%	94%	17%	44%	77%	91%	5%
11	25	# counties	393	71	116	130	11	7	44	12	3
		population	147,943	39,345	26,547	38,494	2,164	4,317	15,739	21,306	173
		% # counties	75%	82%	77%	97%	25%	33%	70%	75%	25%
		% population	81%	92%	65%	100%	29%	46%	85%	94%	8%
10	30	# counties	391	62	133	128	14	7	34	12	2
		population	146,671	35,932	29,392	38,335	2,611	4,856	14,266	21,306	114
		% # counties	74%	71%	89%	96%	32%	33%	54%	75%	17%
		% population	81%	84%	72%	99%	35%	52%	77%	94%	5%
10	25	# counties	417	71	133	130	17	8	44	12	3
		population	151,975	39,345	29,392	38,494	2,782	4,887	15,739	21,306	173
		% # counties	79%	82%	89%	97%	39%	38%	70%	75%	25%
		% population	83%	92%	72%	100%	38%	52%	85%	94%	8%