Appendix G-1

Conceptual Model

The Nature of the Fine Particle and Regional Haze Air Quality Problems in the MANE-VU Region: A Conceptual Description

Prepared for the Ozone Transport Commission

Prepared by NESCAUM
Boston, MA

Final
November 2, 2006

Contributing Authors

Tom Downs, Maine DEP
Iyad Kheirbek, NESCAUM
Gary Kleiman, NESCAUM
Paul Miller, NESCAUM
Leah Weiss, NESCAUM

Acknowledgements

NESCAUM thanks the U.S. EPA's support of the Mid-Atlantic/Northeast Visibility Union Regional Organization whose work has provided the foundational basis of this report.

NESCAUM also thanks the following people for their comments and input during the development of this report:

John Graham, NESCAUM
Kurt Kebschull, Connecticut Department of Environmental Protection
Tonalee Key, New Jersey Department of Environmental Protection
Mohammed A. Majeed, Delaware Department of Natural Resources and Environmental Conservation
David Wackter, Connecticut Department of Environmental Protection
Susan Wierman, MARAMA
Mike Woodman, Maryland Department of the Environment
Martha Webster, Maine Department of Environmental Protection
Jung-Hun Woo, NESCAUM

TABLE OF CONTENTS

FIGURES

TABLES

Executive Summary

Scientific evidence has established a solid link between cardiac and respiratory health risks and transient exposure to ambient fine particle pollution. The same fine particles that are capable of penetrating deep into the lungs are also in the size range that is most efficient at absorbing and scattering visible light, thus impairing visibility. The emission sources, atmospheric chemistry, and meteorological phenomena that influence ambient concentrations of fine particle pollution can act on scales that range from hundreds to thousands of kilometers. Fine particles are not exclusively a secondary pollutant; primary fine particle pollution from local sources can have a significant effect on ambient concentrations in some locations. Fine particles are also not exclusively a summertime pollutant. There are important differences between the meteorological and chemical dynamics that are responsible for high fine particle levels during summer and winter.

In 1997, the U.S. Environmental Protection Agency (USEPA) issued a national ambient air quality standard (NAAQS) for fine particles with an aerodynamic diameter of 2.5 micrometers or less. In 1999, the USEPA followed up with the Regional Haze Rule that enforces a national visibility goal laid out in the Clean Air Act. This will ultimately restore natural visibility to 156 national parks and wilderness areas across the country (called "Class I" areas). To address these Clean Air Act requirements, states will have to develop State Implementation Plans (SIPs) detailing their approaches for reducing fine particle pollution to meet the health-based fine particle NAAQS. They also must develop plans that address the degradation of visibility that exists in various parts of the Northeast (referred to as the Mid-Atlantic/Northeast Visibility Union (MANE-VU) region). As part of this process, the USEPA urges states to include in their SIPs a conceptual description of the pollution problem in their nonattainment and Class I areas. This document provides the conceptual description of the fine particulate and regional haze problems in the MANE-VU states consistent with the USEPA's guidance.

Scientific studies of the regional fine particle problem have uncovered a rich complexity in the interaction of meteorology and topography with fine particle formation and transport. Large scale high pressure systems covering hundreds of thousands of square miles are the source of classic severe fine particle episodes in the eastern United States, particularly in summer. These large, synoptic scale systems create particularly favorable conditions for the oxidation of sulfur dioxide (SO_2) emissions to various forms of sulfate which, in turn, serves to form – or is incorporated into – fine particles that are subsequently transported over large distances. These synoptic scale systems move from west to east across the United States, bringing air pollution emitted by large coal-fired power plants and other sources located outside MANE-VU into the region. This then adds to the pollution burden within MANE-VU on days when MANE-VU's own air pollution sources are themselves contributing to poor air quality. At times, the high pressure systems may stall over the East for days, creating particularly intense fine particle episodes.

In the winter, temperature inversions occur that are effective at concentrating local primary particle emissions at the surface overnight and during early morning hours. This pollution can then be mixed into regionally transported particle pollution (aloft) later

in the morning when convection is restored. Additionally, the lower temperature in the winter can shift the chemical equilibrium in the atmosphere slightly toward the production of nitrate particle pollution relative to sulfate formation. As a result, nitrate can become a significant fraction of measured fine particle mass in parts of the eastern U.S. during winter months.

Primary and secondary emissions of carbon-containing compounds (e.g., diesel exhaust, biogenic organic carbon emissions, and anthropogenic volatile organic compound emissions) all contribute to a significant presence of carbonaceous aerosol across the MANE-VU region, which can vary from urban to rural locations and on a seasonal basis. In addition, short range pollution transport exists, with primary and precursor particle pollutants pushed by land, sea, mountain, and valley breezes that can selectively affect relatively local areas. With the knowledge of the different emission sources, transport scales, and seasonal meteorology in various locations adjacent to and within MANE-VU, a conceptual picture of fine particle pollution and its impacts emerges.

The conceptual description that explains elevated regional $PM_{2.5}$ peak concentrations in the summer differs significantly from that which explains the largely urban peaks observed during winter. On average, summertime concentrations of sulfate in the northeastern United States are more than twice that of the next most important fine particle constituent, organic carbon (OC), and more than four times the combined concentration of nitrate and black carbon (BC) constituents. Episodes of high summertime sulfate concentrations are consistent with stagnant meteorological flow conditions upwind of the MANE-VU region and the accumulation of airborne sulfate (via atmospheric oxidation of SO_2) followed by long-range transport of sulfur emissions from industrialized areas within and outside the region.

National assessments have indicated that in the winter, sulfate levels in urban areas are higher than background sulfate levels across the eastern U.S., indicating that the local urban contribution to wintertime sulfate levels is significant relative to the regional sulfate contribution from long-range transport. A network analysis for the winter of 2002 suggests that the local enhancement of sulfate in urban areas of the MANE-VU region ranges from 25 to 40% and that the long-range transport component of $PM_{2.5}$ sulfate is still the dominant contributor in most eastern cities.

In the winter, urban OC and sulfate each account for about a third of the overall $PM_{2.5}$ mass concentration observed in Philadelphia and New York City. Nitrate also makes a significant contribution to urban $PM_{2.5}$ levels observed in the northeastern United States during the winter months. Wintertime concentrations of OC and nitrate in urban areas can be twice the average regional concentrations of these pollutants, indicating the importance of local source contributions. This is likely because winter conditions are more conducive to the formation of local inversion layers which prevent vertical mixing. Under these conditions, emissions from tailpipe, industrial and other local sources become concentrated near the Earth's surface, adding to background pollution levels associated with regionally transported emissions.

From this conceptual description of fine particle pollution formation and transport into and within MANE-VU, air quality planners need to develop an understanding of

what it will take to clean the air in the MANE-VU region. Every air pollution episode is unique in its specific details. The relative influences of the transport pathways and local emissions vary by hour, day, and season. The smaller scale weather patterns that affect pollution accumulation and its transport underscore the importance of local (in-state) controls for SO_2, nitrogen oxides (NO_X) and volatile organic compound (VOC) emissions. Larger synoptic scale weather patterns, and pollution patterns associated with them, support the need for SO_2 and NO_X controls across the broader eastern United States. Studies and characterizations of nocturnal low level jets also support the need for local and regional controls on SO_2 and NO_X sources as locally generated and transported pollution can both be entrained in low level jets formed during nighttime hours. The presence of land, sea, mountain, and valley breezes indicate that there are unique aspects of pollution accumulation and transport that are area-specific and will warrant policy responses at the local and regional levels beyond a one-size-fits-all approach.

The mix of emission controls is also important. Regional fine particle formation is primarily due to SO_2, but NO_X is also important because of its influence on the chemical equilibrium between sulfate and nitrate pollution during winter. While the effect of reductions in anthropogenic VOCs is less well characterized at this time, secondary organic aerosol (SOA) is a major component of fine particles in the region and reductions in anthropogenic sources of OC may have a significant effect on fine particle levels in urban nonattainment areas. Therefore, a combination of localized NO_X and VOC reductions in urban centers with additional SO_2 and NO_X reductions from across a larger region will help to reduce fine particles and precursor pollutants in nonattainment areas as well improve visibility across the entire MANE-VU region.

1. INTRODUCTION

1.1. Background

Fine particle pollution is a persistent public health problem in the Mid-Atlantic/Northeast Visibility Union (MANE-VU) region. Because of its physical structure, fine particulate matter ($PM_{2.5}$) can bypass conductive airways and deliver exogenous materials, such as reactive organic chemicals that adsorb onto the particle core, into the deep lung.[1] Studies of particulate matter (PM) in urban areas have found associations of short- (daily) and long-term (annual and multiyear) exposure to airborne PM as well as $PM_{2.5}$ with cardiopulmonary health outcomes. These effects include increased symptoms, hospital admissions and emergency room visits, and premature death (Pope *et al.* 2004).

In addition to health implications, visibility impairment in the eastern United States is largely due to the presence of light-absorbing and light-scattering fine particles in the atmosphere. The United States Environmental Protection Agency (USEPA) has identified visibility impairment as the best understood of all environmental effects of air pollution (Watson, 2002). A long-established physical and chemical theory relates the interaction of particles and gases in the atmosphere with the transmission of visual information along a sight path from object to observer.

The Clean Air Act requires states that have areas designated "nonattainment" of the fine particle national ambient air quality standard (NAAQS) to submit State Implementation Plans (SIPs) demonstrating how they plan to attain the fine particle NAAQS.[2] The Clean Air Act also contains provisions for the restoration and maintenance of visibility in 156 federal Class I areas.[3] SIPs for dealing with visibility impairment (or regional haze) must include a long-term emissions management strategy aimed at reducing fine particle pollution in these rural areas.

As part of the SIP process for both of these air quality issues, the USEPA urges states to include a conceptual description of the pollution problem. The USEPA has provided guidance on developing a conceptual description, which is contained in Chapter 11 of the document "Guidance on the Use of Models and Other Analyses for

---

[1] $PM_{2.5}$ or "fine particles" refer to those particles with a diameter ≤ 2.5 micrometers (μm).

[2] The 1997 $PM_{2.5}$ NAAQS includes a requirement that the three-year average of yearly annual average $PM_{2.5}$ design values must be below 15 $\mu g/m^3$ and a requirement that the three-year average of the 98th percentile 24-hour average concentration must be below 65 $\mu g/m^3$. In October 2006, the USEPA acted to change the daily standard (98th percentile value based on valid 24-hour average concentrations measured at a site) from 65 to 35 $\mu g/m^3$.

[3] The Class I designation applies to national parks exceeding 6,000 acres, wilderness areas and national memorial parks exceeding 5,000 acres, and all international parks that were in existence prior to 1977. In the MANE-VU area, this includes: Acadia National Park, Maine; Brigantine Wilderness (within the Edwin B. Forsythe National Wildlife Refuge), New Jersey; Great Gulf Wilderness, New Hampshire; Lye Brook Wilderness, Vermont; Moosehorn Wilderness (within the Moosehorn National Wildlife Refuge), Maine; Presidential Range – Dry River Wilderness, New Hampshire; and Roosevelt Campobello International Park, New Brunswick.

Demonstrating Attainment of Air Quality Goals for Ozone, $PM_{2.5}$, and Regional Haze" (EPA-Draft 3.2, September 2006) (Appendix A of this report reproduces Chapter 11 of the USEPA guidance document). This report provides the MANE-VU states with the basis for their conceptual descriptions, consistent with the USEPA's guidance. In the guidance, the USEPA recommends addressing 13 questions related to $PM_{2.5}$ and eight questions related to visibility to help define the problem in a nonattainment or Class I area. This report addresses these questions, as well as provides some in-depth data and analyses that can assist states in developing conceptual descriptions tailored to their specific areas.

1.2. PM Formation

Fine particles directly emitted into the atmosphere are called "primary" fine particles, and they come from both natural and human sources. These fine particles commonly include unburned carbon particles directly emitted from high-energy processes such as combustion, and particles emitted as combustion-related vapors that condense within seconds of being exhausted to ambient air. Combustion sources include motor vehicles, power generation facilities, industrial facilities, residential wood burning, agricultural burning, and forest fires.

Fine particles are also comprised of "secondary" fine particles, which are formed from precursor gases reacting in the atmosphere or through the addition of PM to pre-existing particles. Although direct nucleation from the gas phase is a contributing factor, most secondary material accumulates on pre-existing particles in the 0.1 to 1.0 micrometer (µm) range and typically account for a significant fraction of the fine PM mass. Examples of secondary particle formation include the conversion of sulfur dioxide (SO_2) to sulfuric acid (H_2SO_4) droplets that further react with ammonia (NH_3) to form various sulfate particles (e.g., ammonium sulfate $(NH_4)_2SO_4$, ammonium bisulfate (NH_4HSO_4), and letovicite ($(NH_4)_3H(SO_4)_2$). The dominant source of SO_2 emissions in the eastern U.S. is fossil fuel combustion, primarily at coal-fired power plants and industrial boilers. Similarly, secondary $PM_{2.5}$ is created by the conversion of nitrogen dioxide (NO_2) to nitric acid (HNO_3) which reacts further with ammonia to form ammonium nitrate (NH_4NO_3) particles. Nitrate particles are formed from the NO_X emitted by power plants, automobiles, industrial boilers, and other combustion sources. Nitrate production in the northeastern U.S. is ammonia-limited and controlled by the availability of sulfate and temperature, especially along the East Coast.[4] While human sources account for most nitrate precursors in the atmosphere, there are some natural sources, including lightning, biological and abiological processes in soils, and stratospheric intrusion. Large sources of ammonia arise from major livestock production and fertilizer application throughout the Midwest, Gulf Coast, mid-Atlantic, and southeastern United States, in addition to the sources of ammonia associated with human activities.

The carbon fraction of fine PM may refer to black carbon (BC) and primary organic and/or secondary organic carbon (OC). Most black carbon is primary, which is

[4] Ammonia reacts preferentially with sulfuric acid, and if sufficient excess ammonia is available, it can then combine with nitric acid to form particulate nitrate.

also sometimes referred to as elemental carbon (EC) or soot. Black carbon is the light-absorbing carbonaceous material in atmospheric particles caused by the combustion of diesel, wood, and other fuels. Organic carbon includes both primary emissions and secondary organic PM in the atmosphere. Secondary organic particles are formed by reactions involving volatile organic compounds (VOCs), which yield compounds with low saturation vapor pressures that nucleate or condense on existing particles at ambient temperature. Organic carbon in both the gas and solid phase is emitted by automobiles, trucks, and industrial processes, as well as by many types of vegetation. The relative amounts of organic carbon from different sources remain highly uncertain, and data are needed to be able to assess the relative contribution of primary versus secondary and anthropogenic versus biogenic production.

1.3. PM Impacts on Visibility

Under natural atmospheric conditions, the view in the eastern United States would extend about 60 to 80 miles (100 to 130 kilometers) (Malm, 2000). Unfortunately, views of such clarity have become a rare occurrence in the East. As a result of man-made pollution, the average visual range in the eastern half of the country has diminished to about 15-30 miles, approximately one-third the visual range that would be observed under unpolluted natural conditions.

In general, the ability to see distant features in a scenic vista is determined less by the amount of light reaching the observer than by the contrast between those features and their surroundings. For example, the illumination of a light bulb in a greenhouse is barely discernible on a sunny day but would be highly visible at night. Similarly, a mountain peak is easily seen if it appears relatively dark against the sunlit sky. If, on the other hand, a milky haze "fills" the space between the observer and the mountain peak, the contrast between the mountain and its background is diminished as both take on a similar hue (Figure 1-1).

Figure 1-1. View of a good visibility day (left) and a poor visibility day (right) at Acadia National Park, Maine in June 2003.

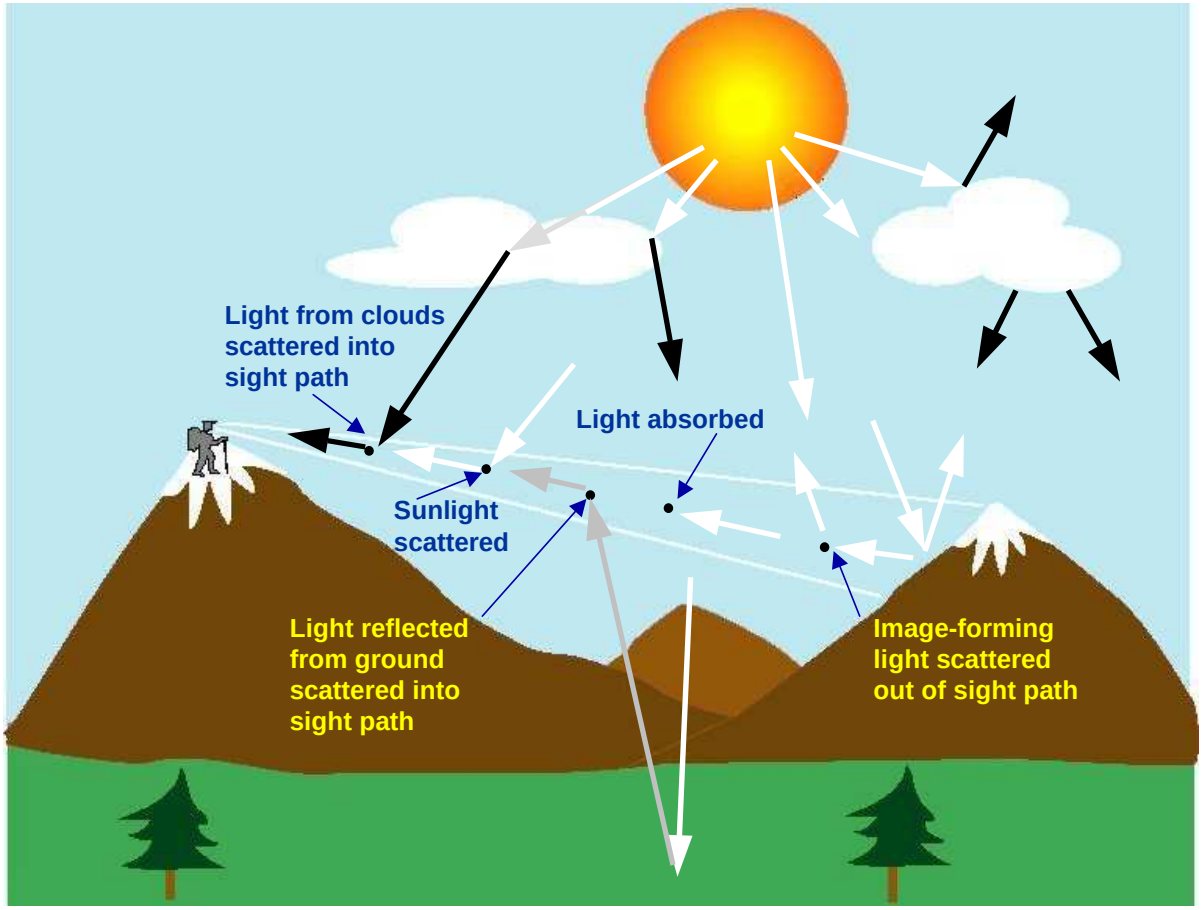
Source: CAMNET, http://www.hazecam.net

In simple terms, this hazy effect occurs when small particles and certain gaseous molecules in the atmosphere absorb or scatter visible light, thereby reducing the amount of visual "information" that reaches the observer. This occurs to some extent even under natural conditions, primarily as a result of the light scattering effect of individual air

molecules (known as Rayleigh scattering[5]) and of naturally occurring aerosols.[6] The substantial visibility impairment caused by manmade pollution, however, is almost entirely attributable to the increased presence of fine particles in the atmosphere.[7]

Figure 1-2 presents a simplified schematic of the way such small particles interact with packets of light or "photons" as they travel from a distant object to an observer. Along the way, particles suspended in the air can deflect or scatter some of the photons out of the sight path. Intervening particles can also absorb photons, similarly removing them from the total amount of light reaching the observer.

Figure 1-2. Schematic of visibility impairment due to light scattering and absorption (adapted from Malm, 2000).



[5] Because air molecules more effectively scatter light of short wavelengths (i.e., blue light), Rayleigh scattering explains the blue color of the sky.

[6] Atmospheric aerosol is a more general term for fine particles suspended in the atmosphere and refers to any particle (solid or liquid) that is suspended in the atmosphere.

[7] The only light-absorbing *gaseous* pollutant present in the atmosphere at significant concentrations is nitrogen dioxide (NO_2). However, the contribution of NO_2 to overall visibility impacts in the Northeast is negligible and hence its effects are not generally included in this discussion or in standard calculations of visibility impairment.

At the same time, particles in the air can scatter light into the sight path, further diminishing the quality of the view. The extraneous light can include direct sunlight and light reflected off the ground or from clouds. Because it is not coming directly from the scenic element, this light contains no visual information about that element. When the combination of light absorption and light scattering (both into and out of the sight path) occurs in many directions due to the ubiquitous presence of small particles in the atmosphere, the result is commonly described as "haze."

1.4. $PM_{2.5}$ Design Values in the MANE-VU Region

SIP developers use monitoring data in several important ways to support SIP activities. This section as well as Section 1.5 present measurements from the FRM and IMPROVE network needed in establishing SIP requirements. Following USEPA guidance (40CFR Part 50, Appendix N; USEPA, 2003a; USEPA, 2003b), we use these data to preview the Design Values and Baseline Conditions that SIP developers must consider for each nonattainment area and Class I area.

The current annual fine particle National Ambient Air Quality Standard was established in 1997 at 15 $\mu g/m^3$. To meet this standard, the 3-year average of a site's annual mean concentration must not be greater than this level. The current daily standard was set at 65 $\mu g/m^3$ at the 98^{th} percentile level. To meet this standard, the 98^{th} percentile value (of valid measurements recorded at a site) must not be greater than this level. No counties in MANE-VU have been designated nonattainment for the daily standard, however, the USEPA has revised the NAAQS with respect to the 24-hr average concentrations and states will have to comply with the new standard (35 $\mu g/m^3$ at the 98^{th} percentile level) within five years of designations (expected in 2010). Fine particle data from the USEPA's Air Quality System (AQS) database for years 2002 through 2004 were used to determine the attainment status of monitoring sites in MANE-VU.

Table 1-1 shows a summary of areas found to exceed the annual standard (no areas exceed the daily standard). As tabulated, 12 areas fail to achieve the annual standard, with design values ranging from 15.1 to 20.4 $\mu g/m^3$. The nonattainment areas are concentrated in Pennsylvania and the coastal urban corridor. Sulfates and organic carbon represent the largest contributors to these high fine particle levels.

Table 1-1. 2004 $PM_{2.5}$ Design Value for Nonattainment Areas in MANE-VU

State(s)	Nonattainment Area	2004 Annual Design Value	2004 24-hr Design Value
MD	Baltimore	16.3	41
PA	Harrisburg-Lebanon-Carlisle	15.4	41
PA	Johnstown	15.3	40
PA	Lancaster	16.8	42
PA	Liberty-Clairton	20.4	65
MD	Martinsburg, WV-Hagerstown	16.1	39
NY-NJ-CT	New York-N. New Jersey-Long Island	16.8	50
PA-NJ-DE	Philadelphia-Wilmington	15.4	39
PA	Pittsburgh-Beaver Valley	16.5	45
PA	Reading	16.1	42
DC-MD-VA	Washington, DC	15.1	42
PA	York	16.9	43

1.5. Regional haze baseline conditions

The Regional Haze Rule requires states and tribes to submit plans that include calculations of current and estimated baseline and natural visibility conditions. They will use monitoring data from the IMPROVE program as the basis for these calculations. Table 1-2 and Table 1-3 present the five-year average[8] of the 20 percent worst day mass concentrations and 20 percent best day mass concentrations respectively in six Class I areas. Five of these areas are in MANE-VU and one (Shenandoah) is nearby but located in a neighboring regional planning organization (RPO) region.[9] Table 1-4 and Table 1-5 give the corresponding worst day and best day contributions to particle extinction for the six Class I areas. Each of these tables show the relative percent contribution for all six Class I sites. Sulfate and organic carbon dominate the fine mass, with sulfate even more important to particle extinction.

To guide the states in calculating baseline values of reconstructed extinction and for estimating natural visibility conditions, the USEPA released two documents in the fall of 2003 outlining recommended procedures (USEPA 2003a; USEPA 2003b). Recently, the IMPROVE Steering Committee endorsed an alternative method for the calculation of these values. The IMPROVE alternative methods were used, to create Table 1-6, which provides detail on the uniform visibility goals for the 20 percent worst conditions at the six Class 1 areas.

[8] Great Gulf calculations are based on four years of data (2001-2004).

[9] Note that values presented for Shenandoah, a Class I area in the Visibility Improvement State and Tribal Association of the Southeast (VISTAS) region, are for comparative purposes only. VISTAS will determine uniform rates of progress for areas within its region.

The first column of data in Table 1-6 gives the alternative proposed natural background levels for the worst visibility days at these six sites. MANE-VU has decided to use this approach, at least initially, for 2008 SIP planning purposes (NESCAUM, 2006). The second column shows the baseline visibility conditions on the 20 percent worst visibility days. These values are based on IMPROVE data from the official five-year baseline period (2000-2004) and again were calculated using the IMPROVE alternative approach. Using these baseline and natural background estimates, we derive the uniform rate of progress shown in the third column.[10] The final column displays the interim 2018 progress goal based on 14 years of improvement at the uniform rate.

Table 1-2. Fine mass and percent contribution for 20 percent worst days

20% Worst-day Fine Mass (μg/m^3)/% contribution to fine mass					
Site	SO_4	NO_3	OC	EC	Soil
Acadia	6.3/ 56%	0.8/ 7%	3.2/ 28%	0.4/ 4%	0.5/ 5%
Brigantine	11.6/ 56%	1.7/ 8%	5.8/ 28%	0.7/ 3%	1/ 5%
Great Gulf	7.3/ 59%	0.4/ 3%	3.8/ 31%	0.4/ 3%	0.6/ 5%
Lye Brook	8.5/ 58%	1.1/ 7%	3.9/ 27%	0.5/ 3%	0.6/ 4%
Moosehorn	5.7/ 54%	0.7/ 7%	3.4/ 32%	0.4/ 4%	0.4/ 4%
Shenandoah	13.2/ 68%	0.7/ 3%	4.2/ 22%	0.6/ 3%	0.7/ 4%

Table 1-3. Fine mass and percent contribution for 20 percent best days

20% Best-day Fine Mass (μg/m^3)/% contribution to fine mass					
Site	SO_4	NO_3	OC	EC	Soil
Acadia	0.8/ 42%	0.1/ 6%	0.8/ 41%	0.1/ 5%	0.1/ 6%
Brigantine	1.8/ 43%	0.5/ 11%	1.5/ 35%	0.2/ 6%	0.2/ 5%
Great Gulf	0.7/ 43%	0.1/ 7%	0.7/ 40%	0.1/ 5%	0.1/ 6%
Lye Brook	0.6/ 44%	0.1/ 11%	0.4/ 33%	0.1/ 5%	0.1/ 7%
Moosehorn	0.8/ 37%	0.1/ 6%	1/ 47%	0.1/ 5%	0.1/ 5%
Shenandoah	1.4/ 45%	0.5/ 16%	1/ 29%	0.2/ 5%	0.2/ 5%

[10] We calculate the rate of progress as (baseline – natural background)/60 to yield the annual deciview (dv) improvement needed to reach natural background conditions in 2064, starting from the 2004 baseline.

Table 1-4. Particle extinction and percent contribution for 20 percent worst days

20% Worst-day particle extinction (Mm^{-1}) /% Contribution to particle extinction						
Site	SO_4	NO_3	OC	EC	Soil	CM
Acadia	69.2/ 64%	8/ 7%	11.2/ 10%	4.3/ 4%	0.5/ 0%	1.9/ 2%
Brigantine	127.1/ 66%	15.7/ 8%	24.2/ 13%	7/ 4%	1/ 1%	5.4/ 3%
Great Gulf	76.6/ 68%	3/ 3%	14.4/ 13%	3.9/ 3%	0.6/ 1%	3/ 3%
Lye Brook	87.3/ 67%	9.1/ 7%	15.3/ 12%	4.8/ 4%	0.6/ 0%	1.8/ 2%
Moosehorn	58.5/ 60%	6.4/ 7%	11.9/ 12%	4.4/ 5%	0.4/ 0%	2.1/ 3%
Shenandoah	155.5/ 79%	5.8/ 3%	16.1/ 8%	5.7/ 3%	0.7/ 0%	2.5/ 1%

Table 1-5. Particle extinction and percent contribution for 20 percent best days

20% Best-day particle extinction (Mm^{-1}) /% Contribution to particle extinction						
Site	SO_4	NO_3	OC	EC	Soil	CM
Acadia	6.8/ 28%	1.1/ 4%	2.2/ 9%	0.9/ 4%	0.1/ 0%	0.7/ 6%
Brigantine	14.8/ 35%	3.9/ 9%	4.5/ 11%	2.4/ 6%	0.2/ 1%	3.2/ 11%
Great Gulf	5.8/ 27%	1/ 4%	2/ 9%	0.8/ 4%	0.1/ 0%	0.9/ 8%
Lye Brook	4.4/ 23%	1.2/ 6%	1.3/ 7%	0.6/ 3%	0.1/ 0%	0.5/ 6%
Moosehorn	6.7/ 26%	1.1/ 4%	3.1/ 12%	1/ 4%	0.1/ 0%	1.1/ 8%
Shenandoah	11.2/ 36%	4.2/ 13%	2.9/ 9%	1.6/ 5%	0.2/ 1%	1.1/ 5%

Table 1-6. Natural background and baseline calculations for select Class I areas

Site	**20 % Worst Days Natural Background (dv)**	**20% Worst Days Baseline 2000-04(dv)**	**Uniform Rate (dv/yr)**	**Interim Progress Goal 2018 (dv)**	**20% Best Days Baseline 2000-04(dv)**
Acadia	12.54	22.89	0.17	20.47	8.77
Brigantine	12.34	29.01	0.28	25.12	14.33
Great Gulf	12.12	22.82	0.18	20.32	7.66
Lye Brook	11.85	24.44	0.21	21.50	6.37
Moosehorn	12.10	21.72	0.16	19.48	9.15
Dolly Sods	10.45	29.05	0.31	24.71	12.28
James River Face	11.20	29.12	0.30	24.94	14.21
Shenandoah	11.44	29.31	0.30	25.14	10.92

As demonstrated in Table 1-2, the inorganic constituents of fine particles, sulfates and nitrates are the dominant contributors to visibility impairment, accounting for about 80 percent of total particle extinction. Within the MANE-VU sites, the relative split between these two components is ~8 to 1 sulfate to nitrate (at Shenandoah, the average 20 percent worst day contribution of sulfates is even more dominant). Carbonaceous components account for the bulk of the remaining particle extinction, ranging from 12 to nearly 20 percent, mostly in the form of organic carbon. The remaining components add little to the extinction budget on the worst days, with a few percent attributable to coarse mass and around a half percent from fine soil.

References

NESCAUM, *Baseline and Natural Background Visibility Conditions: Considerations and Proposed Approach to the calculation of Baseline and Natural Background Visibility Conditions at MANE-VU Class I areas*, Northeast States for Coordinated Air Use Management, Boston, MA, June 2006.

Pope C.A., R.T. Burnett, G.D. Thurston, et al. "Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease." *Circulation* 109:71-7, 2004.

Malm, W.C., *Introduction to Visibility*, Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Fort Collins, CO, 80523, 2000.

Watson, J., "Visibility: Science and Regulation", *JAWMA* 52:628-713, 2002.

USEPA, "Guidance for Tracking Progress under the Regional Haze Rule" EPA-454/B-03-004 September 2003a.

USEPA, "Guidance for Estimating Natural Visibility Conditions under the Regional Haze Program" EPA-454/B-03-005 September 2003b.

USEPA. "Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, $PM_{2.5}$, and Regional Haze." U.S. Environmental Protection Agency Office of Air Quality Planning and Standards Emissions, Monitoring, and Analysis Division. Draft 3.2- September 2006

2. A DETAILED LOOK AT FINE PARTICLE POLLUTION AND REGIONAL HAZE IN THE MANE-VU REGION

Developing a conceptual description of fine particle pollution or regional haze requires combining experience and atmospheric-science expertise with multiple data sources and analysis techniques. This includes measured data on ambient pollutant concentrations as well as emission inventory and meteorological data, chemical transport modeling, and observationally based models (NARSTO, 2003). Here, we begin with a conceptual description based on the existing scientific literature and regional data analyses concerning $PM_{2.5}$ and its effect on visibility. This includes numerous review articles and reports on the subject. Subsequent chapters review monitoring data, emissions inventory information, and modeling results to support the conceptual understanding of regional fine particle pollution presented here.

Most past assessments of fine particle pollution and visibility impairment have tended to be national in scope. For purposes of this discussion, we have selectively reviewed the literature in order to present a distinctly eastern U.S. focus. While we already know much about fine particle pollution and visibility impairment and their causes in the MANE-VU region (see NESCAUM, 2001, 2006; NARSTO, 2003; Watson, 2002), significant gaps in understanding remain with respect to the nitrate and organic component of $PM_{2.5}$. While research continues, we have assembled the relevant information that is available to provide an overview of our current understanding of the regional context for $PM_{2.5}$ nonattainment and visibility impairment in the MANE-VU region.

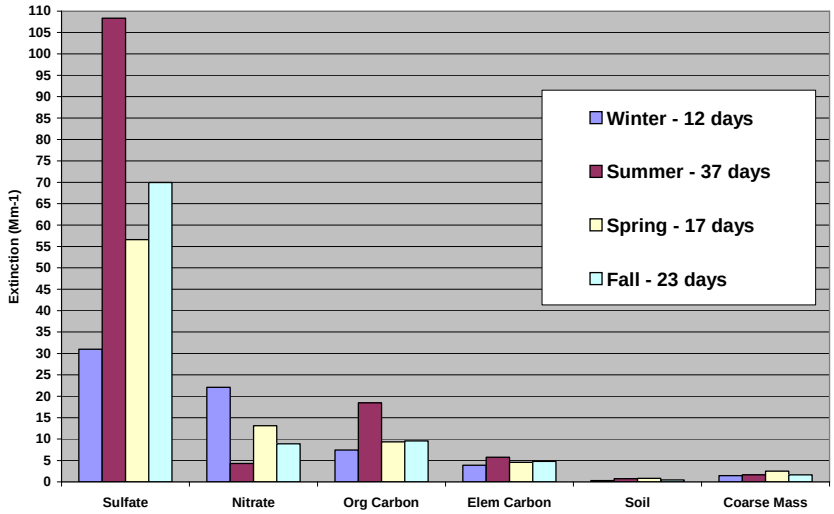
2.1. Chemical composition of particulate matter in the rural MANE-VU region

Sulfate alone accounts for anywhere from one-half to two-thirds of total fine particle mass on high $PM_{2.5}$ days in rural areas of MANE-VU. Even on low $PM_{2.5}$ days, sulfate generally accounts for the largest fraction (40 percent or more) of total fine particle mass in the region (NESCAUM, 2001, 2004b). Sulfate accounts for a major fraction of $PM_{2.5}$, not only in the Northeast but across the eastern United States (NARSTO, 2003).

After sulfate, organic carbon (OC) consistently accounts for the next largest fraction of total fine particle mass. Its contribution typically ranges from 20 to 30 percent of total fine particle mass on the days with the highest levels of $PM_{2.5}$. The fact that the contribution from organic carbon can be as high as 40 percent at the more rural sites on low $PM_{2.5}$ days is likely indicative of the role played by organic emissions from vegetation (so-called “biogenic hydrocarbons”).

Relative contributions to overall fine particle mass from nitrate (NO_3), elemental carbon, and fine soil are all smaller (typically under 10 percent), but the relative ordering among the three species varies with location and season. Figure 2-1 below, reflects the difference between nitrate and organic contributions to rural fine particle concentrations during different seasons (monitoring data for additional sites in the MANE-VU region are in Appendix B).

Figure 2-1. Comparison of contributions during different seasons at Lye Brook Wilderness Area on 20% worst visibility (high $PM_{2.5}$) days (2000-2003).

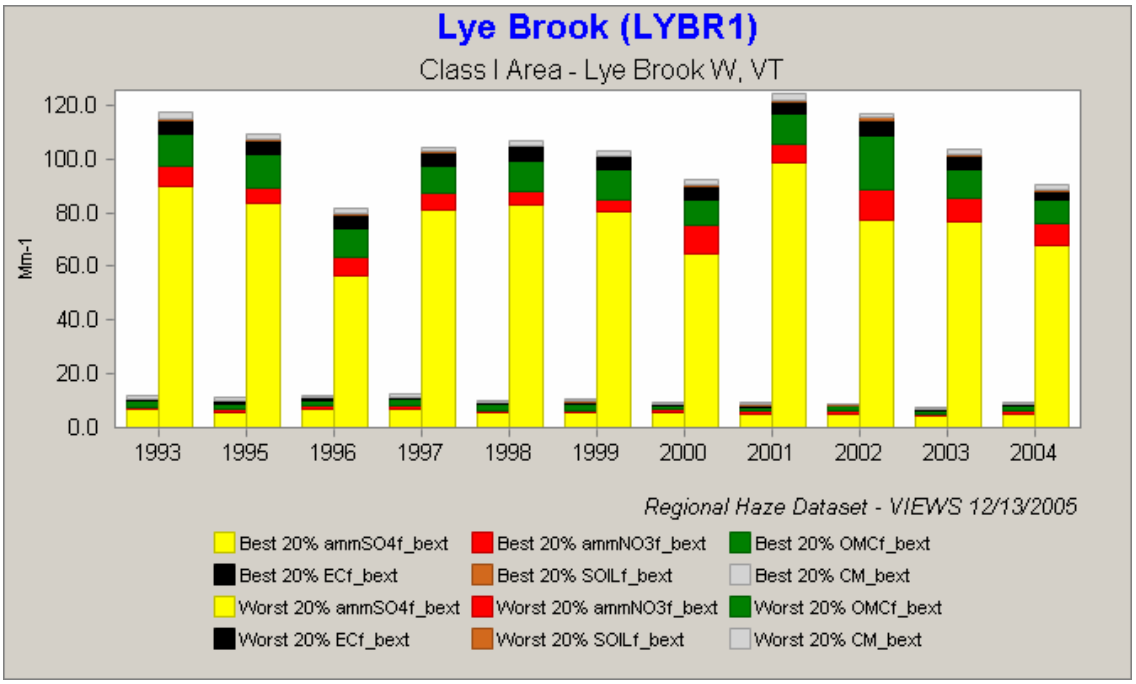


Almost all particle sulfate originates from sulfur dioxide (SO_2) oxidation and typically associates with ammonium (NH_4) in the form of ammonium sulfate ($(NH_4)_2SO_4$). Ninety-five percent of SO_2 emissions are from anthropogenic sources (primarily from fossil fuel combustion), while the majority of ammonium comes from agricultural activities and, to a lesser extent, from transportation sources in some areas (NARSTO, 2003).

Two major chemical pathways produce sulfate from SO_2 in the atmosphere. In the gas phase, production of sulfate involves the oxidation of SO_2 to sulfuric acid (H_2SO_4), ammonium bisulfate (NH_4HSO_4), or ammonium sulfate, depending on the availability of ammonia (NH_3). In the presence of small wet particles (typically much, much smaller than rain drops or even fog), a highly efficient aqueous phase process can oxidize SO_2 to sulfate extremely quickly (~10 percent per hour).

Not only is sulfate the dominant contributor to fine particle mass in the region, it accounts for anywhere from 60 percent to almost 80 percent of the *difference* between fine particle concentrations and extinction on the lowest and highest mass days at rural locations in the northeast and mid-Atlantic states (See Figure 2-2). Notably, at urban locations such as Washington DC, sulfate accounts for only about 40 percent of the difference in average fine particle concentrations for the 20 percent most versus least visibility impaired days (NESCAUM, 2001).

Figure 2-2. Comparison of species contributions on best and worst days at Lye Brook Wilderness Area.



2.2. Rural versus urban chemistry

Contributions to fine particle mass concentrations at rural locations include long-range pollutant transport as well as non-anthropogenic background contributions. Urban areas generally show mean $PM_{2.5}$ levels exceeding those at nearby rural sites. In the Northeast, this difference implies that local urban contributions are roughly 25 percent of the annual mean urban concentrations, with regional aerosol contributing the remaining, and larger, portion (NARSTO, 2003).

This rural versus urban difference in typical concentrations also emerges in a source apportionment analysis of fine particle pollution in Philadelphia (see Chapter 10 of NARSTO, 2003) using two different mathematical models, UNMIX and Positive Matrix Factorization (PMF). This analysis provides additional insight concerning sources of fine particle pollution in urban areas of the densely populated coastal corridor between Washington DC and New England. Specifically, this analysis found the following apportionment of $PM_{2.5}$ mass in the study area:

- Local SO_2 and sulfate: ~ 10 percent
- Regional sulfate: ~ 50 percent
- Residual oil: 4-8 percent
- Soil: 6-7 percent
- Motor vehicles: 25-30 percent

The analysis does not account for biogenic sources, which most likely are embedded in the motor vehicle fraction (NARSTO, 2003). The Philadelphia study suggests that both local pollution from nearby sources and transported "regional"

pollution from distant sources contribute to the high sulfate concentrations observed in urban locations along the East Coast on an annual average basis. Summertime sulfate and organic carbon are strongly regional in eastern North America. Typically 75–95 percent of the urban sulfate concentrations and 60–75 percent of the urban OC concentrations arise from cumulative region-wide contributions (NARSTO, 2003). Urban air pollutants are essentially added on top of this regional background. Nitrate plays a noticeably more important role at urban sites compared to northeastern and mid-Atlantic rural monitoring sites, perhaps reflecting a greater contribution from vehicles and other urban pollution sources (NESCAUM, 2001).

It is difficult to discern any significant meaning about the cause of "excess" mass from a single pair of sites. There are many factors that influence the concentrations at a particular site and it is likely that for every pair of sites that shows an urban excess, one could find some pair of locations that might show something similar to an urban "deficit." While paired sites from an urban and a rural location will *typically* show greater concentrations in the urban location and lower levels of pollution in rural areas, great care must be exercised in the interpretation of any two-site analysis such as the comparisons of speciated components of $PM_{2.5}$ presented here. Nonetheless, such comparisons do provide a general feel for the typical chemical composition of $PM_{2.5}$ in the eastern U.S. and the relative differences in chemical composition between rural and more urban locations. More detailed, "network"-wide analyses (e.g., see NESCAUM 2004b; relevant sections are attached in Appendix C to this report) indicate that the results provided are not anomalous of typical urban environments in the MANE-VU region.

Figure 2-3 and Figure 2-4 compare two urban-rural pairs of speciation monitors: the New York nonattainment area (Elizabeth and Chester, New Jersey) and the Boston metropolitan area (Boston and Quabbin Reservoir, Massachusetts). The first three sites are Speciation Trends locations, while the Reservoir site is part of the IMPROVE protocol network.[11]

[11] To provide a more direct comparison of the differences between the urban and rural sites, only those days for which both monitors in a pair had data were used. Four seasonal averages were computed for 2002, with seasons defined as winter (January, February, December), spring (March, April, May), Summer (June, July, August) and Fall (September, October, November). July 7 was excluded from the analysis because the Quebec forest fires affecting the region on that day would have dominated the summertime averages. The major fine particle species categories considered included ammonium sulfate, ammonium nitrate, organic carbon, elemental carbon, and soil mass. The traditional assumptions about these constituents were made; all sulfate was fully neutralized and a multiplier of 1.4 was used to account for mass of organic carbon. An "other $PM_{2.5}$ mass" category was created to delineate the difference between gravimetric mass determined from the Teflon filter and the reconstructed mass sum of the individual mass constituents. Where no "other" mass is graphed, the sum of the species either equaled or exceeded the directly measured mass. No adjustments were made to account for the different operational definitions of carbon between the IMPROVE and STN networks. Average blank corrections were applied to all samples. In the case of New York City, both rural and urban monitors were STN. The Boston pair reflects not only inter-site differences, but also differences in definition of organic and elemental carbon. However, the general interpretation of the data differences remains consistent. Based on current understanding, the rural elemental carbon would be even lower than what is shown on the graph if it were made consistent with the STN definition of EC. Likewise, the organic carbon value would increase slightly for the rural value, as the

Figure 2-3. New York nonattainment area (Elizabeth, NJ) compared to an upwind background site (Chester, NJ)

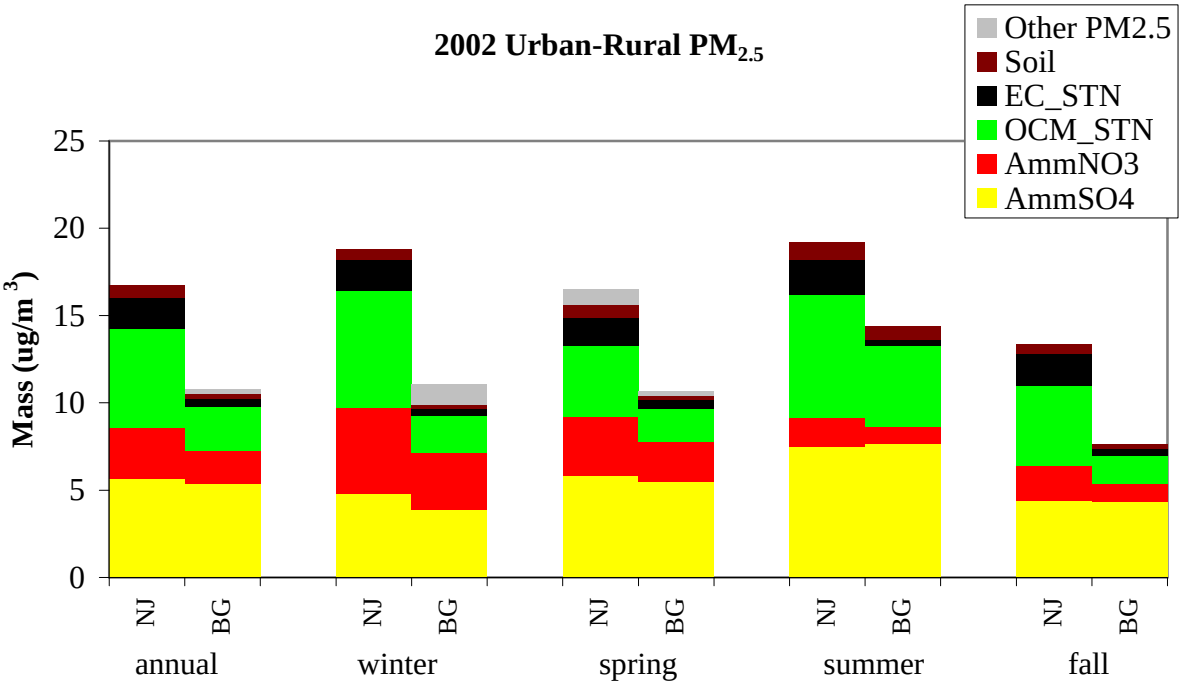
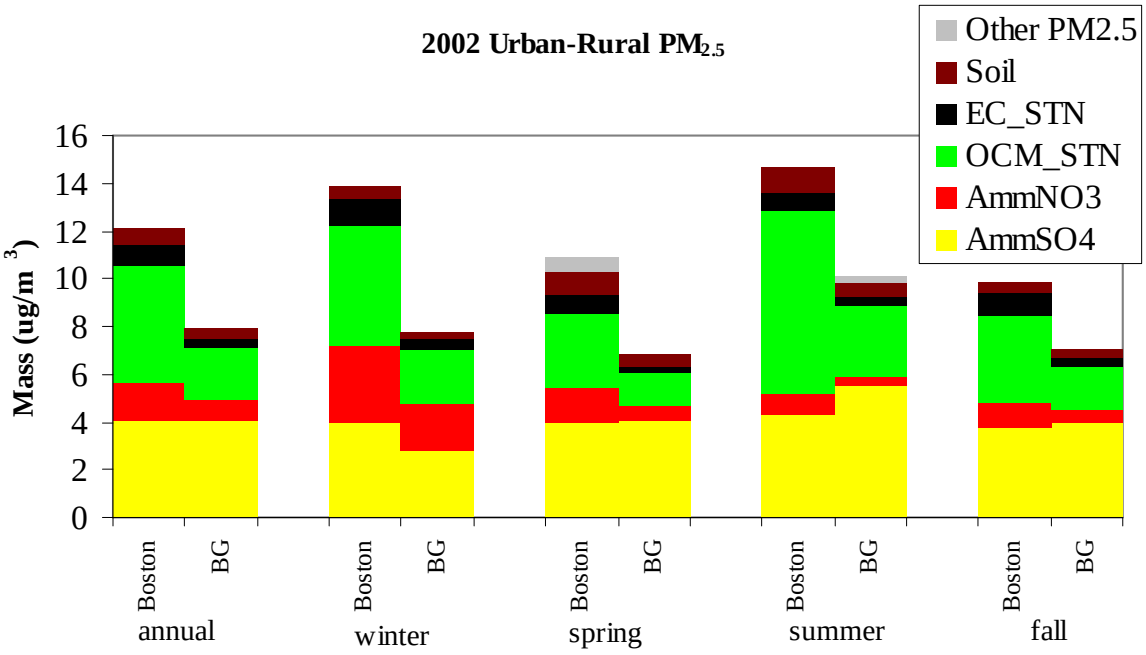


Figure 2-4. Boston urban area (Boston, MA) compared to an upwind background site (Quabbin Reservoir, MA)



EC would be allocated to OC. The urban OC levels are so much greater than those in the rural area that a slight increase in rural OC makes little difference.

The urban-rural differences show consistency for both the New York City nonattainment area and Boston. On an annual scale, the sulfate levels are comparable, with increased mass loading at these urban sites driven primarily by differences in nitrates and carbon with smaller differences in "soil" levels. One interesting aspect of this comparison is the seasonal differences in the urban-rural sulfate split. On an annual basis, sulfate appears to be similar at urban and rural locations (based on these two pair of sites); however, during the colder months, the urban sulfate levels are elevated relative to the rural levels. This behavior is opposite during the summer. During the wintertime, the Northeast urban corridor itself is a substantial source of sulfur emissions. These local emissions can be trapped near the surface during the winter and have a corresponding higher impact on the urban area relative to the rural area.

For both urban and rural areas, the summertime OC levels are significantly greater than wintertime concentrations. Although the oxidation chemistry slows in winter, the cooler temperatures change the phase dynamics, driving more mass into the condensed over the gas phase. This along with more frequent temperature inversions (which limit atmospheric ventilation of the urban boundary layer) can lead to the observed increases in the relative influence of both organic and nitrate levels during winter months. EC, OC, and nitrate all are observed to have higher measured levels in the urban area (but still lower than the comparable summer values measured at the same sites), driven by local sources of these constituents.

2.3. Geographic considerations and attribution of $PM_{2.5}$/haze contributors

In the East, both annual average and maximum daily fine particle concentrations are highest near heavily industrialized areas and population centers. Not surprisingly, given the direct connection between fine particle pollution and haze, the same pattern emerges when one compares measures of light extinction on the most and least visibility impaired days at parks and wilderness areas subject to federal haze regulations in the MANE-VU region (NESCAUM, 2001). An accumulation of particle pollution often results in hazy conditions extending over thousands of square kilometers (km^2) (NARSTO, 2003). Substantial visibility impairment is a frequent occurrence in even the most remote and pristine areas of the MANE-VU region (NESCAUM, 2001).

$PM_{2.5}$ mass declines fairly steadily along a southwest to northeast transect of the MANE-VU region. This decline is consistent with the existence of large fine particle emissions sources (both primary and secondary) to the south and west of MANE-VU. This trend is driven, in large part, by the marked southwest-to-northeast gradient in ambient sulfate concentrations during three seasons of the year as illustrated in Figure 2-5. Wintertime concentrations, by contrast, are far more uniform across the entire region. Figure 2-6 shows that on an annual basis, both total $PM_{2.5}$ and sulfate mass are highest in the southwestern portions of the MANE-VU region (note the different scales for each pollutant). High concentrations of nitrate and organic particle constituents, which play a role in localized wintertime $PM_{2.5}$ episodes, tend to be clustered along the northeastern urban corridor and in other large urban centers.

Figure 2-5. 2002 Seasonal average SO_4 based on IMPROVE and STN data

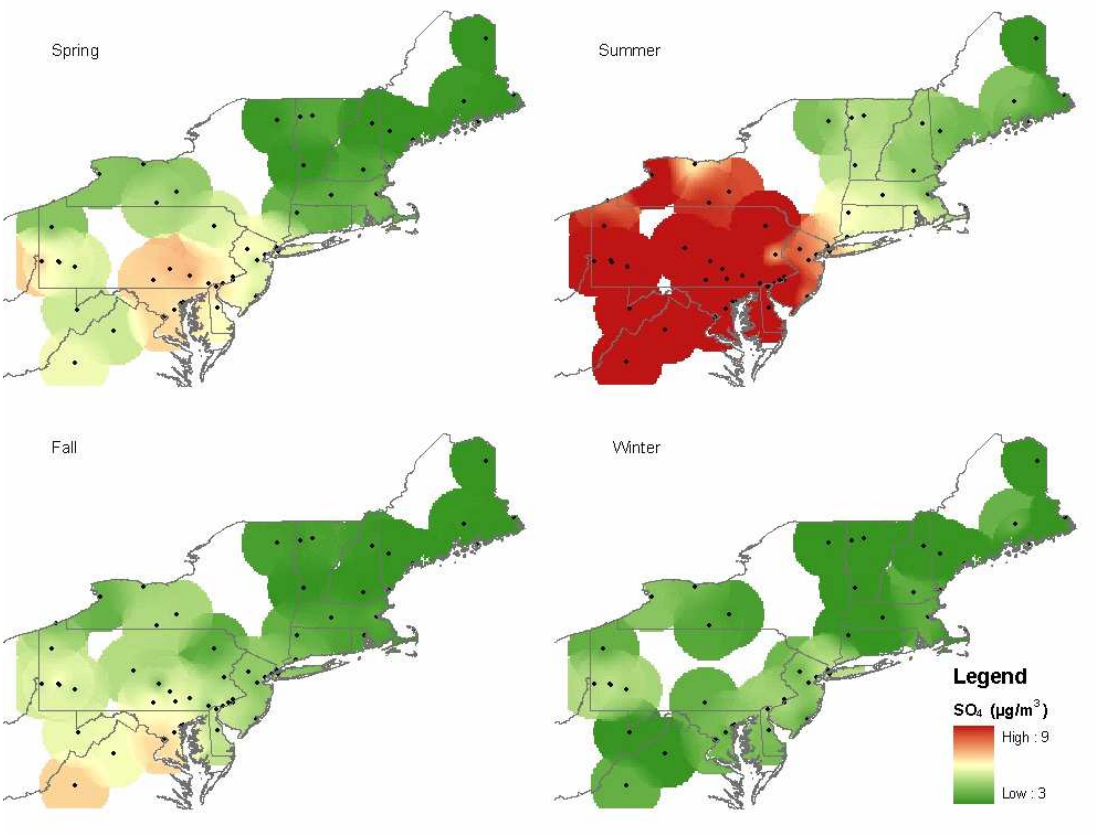
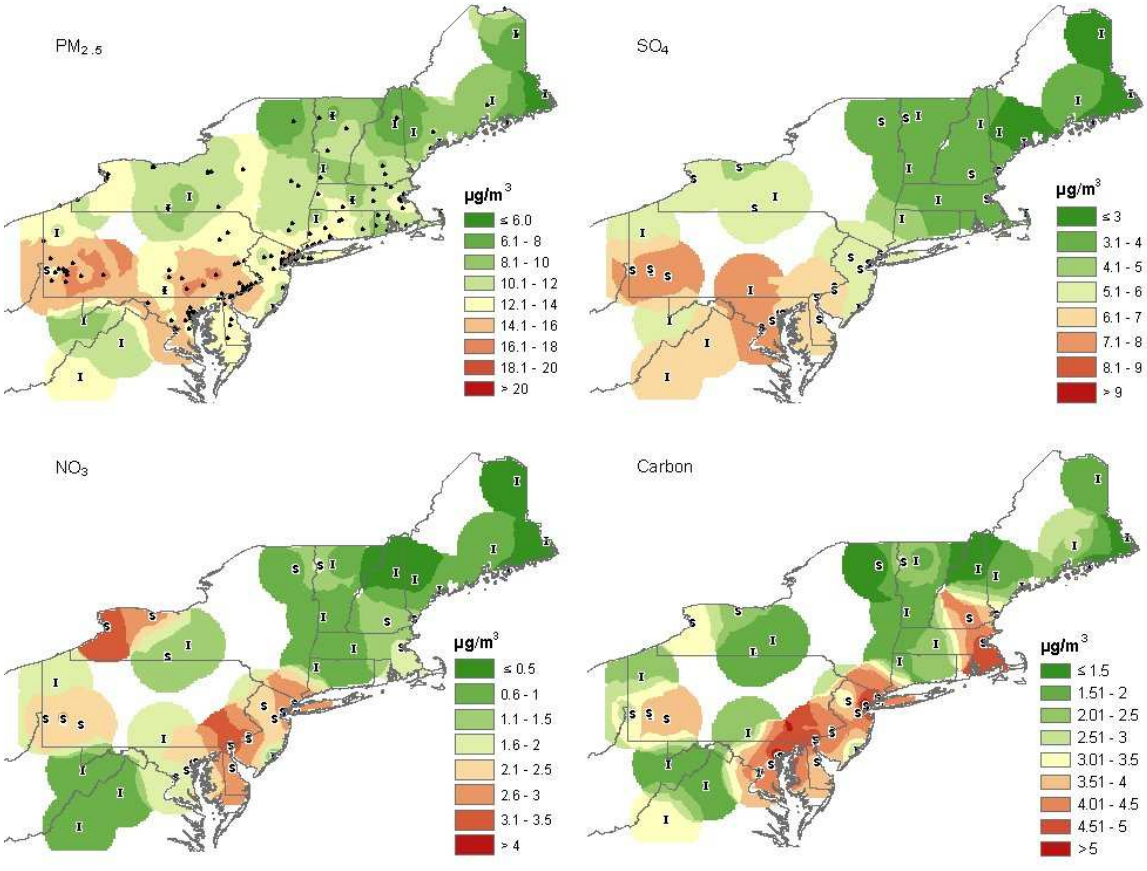


Figure 2-6. 2002 Annual average $PM_{2.5}$, sulfate, nitrate and total carbon for MANE-VU based on IMPROVE (I) and STN (S) data. $PM_{2.5}$ mass data are supplemented by measurements from the FRM network (•).



While these figures provide some preliminary context for identifying sources contributing to the region's particulate matter and visibility problems, they say nothing about the relative efficiency of a state's or region's emissions in contributing to the problem. It is clear that distance from the emissions source matters. Local, nearby sources are exceedingly important and sources within about 200 km are much more efficient (on a per ton emitted basis) at producing pollution impacts at eastern Class I sites such as Shenandoah National Park than emissions sources farther away (USNPS, 2003). In general, the "reach" of sulfate air pollution resulting from SO_2 emissions is longest (650–950 km). The reach of ammonia emissions or reduced nitrogen relative to nutrient deposition is the shortest (around 400 km), while oxides of nitrogen and sulfur — in terms of their impacts with respect to acidic deposition — have a reach between 550–650 km and 600–700 km, respectively (USNPS, 2003).

Monitoring evidence indicates that non-urban visibility impairment in eastern North America is predominantly due to sulfate particles, with organic particles generally second in importance (NARSTO, 2003). This makes sense, given the "long reach" of SO_2 emissions once they are chemically transformed into sulfate and given the ubiquitous nature of OC sources in the East. The poorest visibility conditions occur in highly industrialized areas encompassing and adjacent to the Ohio River and Tennessee Valleys. These areas feature large coal-burning power stations, steel mills, and other large emissions sources. Average fine particle concentrations and visibility conditions are also poor in the highly populated and industrialized mid-Atlantic seaboard but improve gradually northeast of New York City (Watson, 2002).

A review of source apportionment and ensemble trajectory analyses conducted by USEPA (2003) found that all back trajectory analyses for eastern sites associated sulfate with the Ohio River Valley area. These studies also are frequently able to associate other types of industrial pollutants (e.g., copper or zinc smelting, steel production, etc.) with known source areas, lending credibility to their performance. Several studies in the USEPA review noted transport across the Canadian border, specifically sulfates from the midwestern United States into Canada, and smelter emissions from Canada into the northeastern United States.

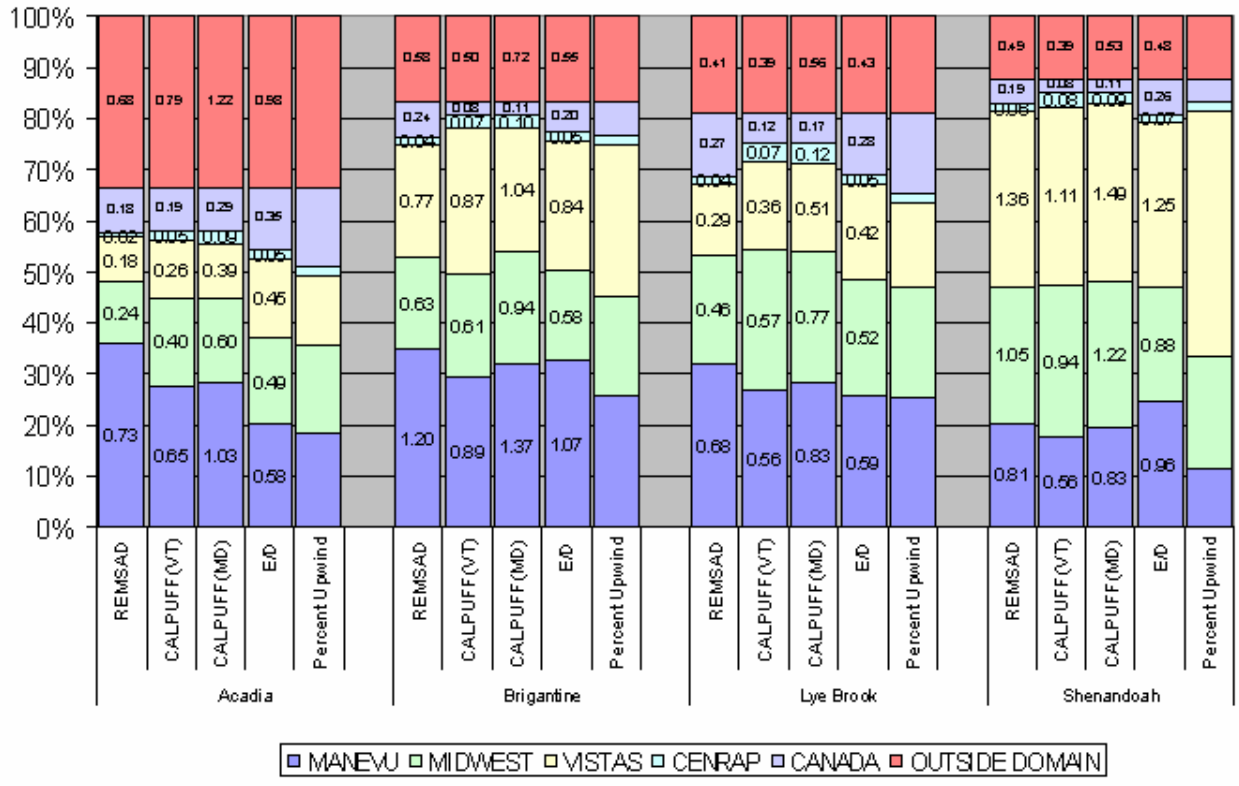
A recent, comprehensive analysis of air quality problems at Shenandoah National Park conducted by the U.S. National Park Service (USNPS, 2003) focused on contributions to particulate pollution and visibility impairment south of the MANE-VU region. In descending order of importance, the Park Service analysis determined that Ohio, Virginia, West Virginia, Pennsylvania, and Kentucky comprise the top five of 13 key states contributing to ambient sulfate concentrations and haze impacts at the park. West Virginia, Ohio, Virginia, Pennsylvania, and Kentucky comprise the top five contributing states with respect to sulfur deposition impacts at the park. Finally, Virginia, West Virginia, Ohio, Pennsylvania, and North Carolina were found to be the top five states contributing to deposition impacts from oxidized nitrogen at the park (USNPS, 2003).

In sum, the Park Service found that emission sources located within a 200 km (125 mile) radius of Shenandoah cause greater visibility and acidic deposition impacts at the park, on a per ton basis, than do more distant emissions sources (USNPS, 2003). When mapping deposition and concentration patterns for all three pollutants using

contour lines, the resulting geographic pattern shows a definite eastward tilt in the area of highest impact. This is the result of prevailing wind patterns, which tend to transport most airborne pollutants in an arc[12] from the north-northeast to the east. The Park Service found, for example, that emissions originating in the Ohio River Valley end up three times farther to the east than to the west (USNPS, 2003).

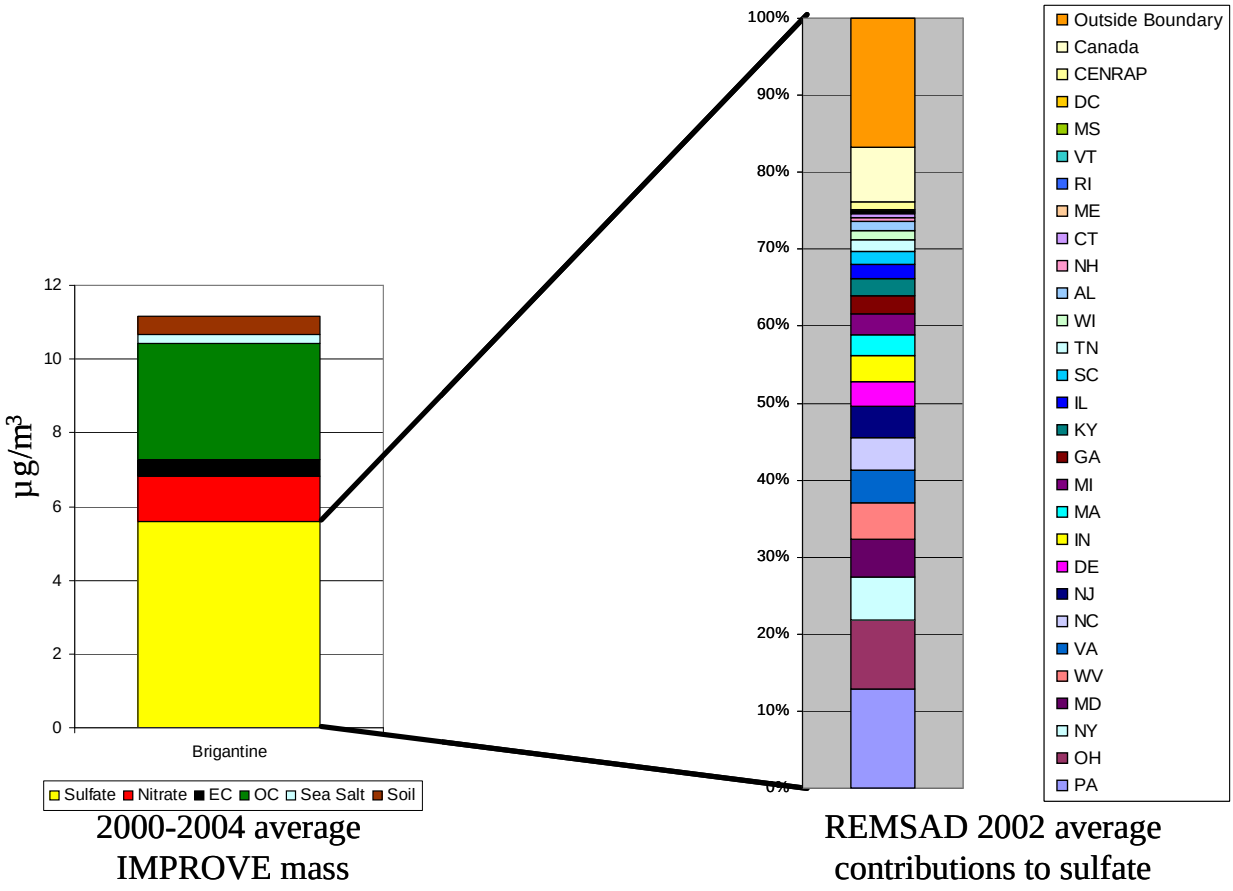
The recent sulfate attribution work completed by MANE-VU (NESCAUM, 2006) finds that a variety of different states contribute to observed sulfate in rural locations across the MANE-VU region, but that in the southwest portions of the region, neighboring RPOs contribute to a more significant degree relative to rural areas in the Northeast. Figure 2-7 shows relative contributions of RPOs to sulfate at three MANE-VU Class I areas and one VISTAS Class I area based on a variety of analysis methods. Figure 2-8 shows the individual state contributions to sulfate at Brigantine Wilderness Area on the New Jersey coast according to tagged REMSAD modeling.

Figure 2-7. 2002 Annual average contribution to $PM_{2.5}$ sulfate as determined by multiple analysis methods for four Class I areas spanning MANE-VU and Virginia



[12] The prevailing winds are eastward to northeast. This leads to greater pollution transport to the east-northeast relative to other directions.

Figure 2-8. 2002 Annual average mass contribution to $PM_{2.5}$ at Brigantine Wilderness in New Jersey (IMPROVE) and sulfate contributions as determined by tagged REMSAD model simulations (NESCAUM, 2006)



2.4. CAIR Modeling

The CAIR modeling by the USEPA provides information on the upwind areas (by state) contributing to downwind nonattainment for $PM_{2.5}$ in MANE-VU counties. Table 2-1 presents the upwind states significantly contributing to $PM_{2.5}$ nonattainment in counties within MANE-VU during 2001, according to significance criteria used by the USEPA (USEPA, 2005, from Table VII-3). The states listed in the table as significantly contributing to downwind nonattainment in MANE-VU counties include states outside of MANE-VU, indicating the broad regional scale of the $PM_{2.5}$ transport problem.

Table 2-2 provides the maximum contribution from each state to annual average $PM_{2.5}$ nonattainment in a downwind state (not necessarily restricted to MANE-VU nonattainment counties) based on CAIR modeling.

Table 2-1. Upwind states that make a significant contribution to $PM_{2.5}$ in each downwind nonattainment county (2001 modeling).

Downwind State/County		Upwind States									
DE	New Castle	MD/DC	MI	NY	OH	PA	VA	WV			
DC	District of Columbia	NC	OH	PA	VA	WV					
MD	Anne Arundel	NC	OH	PA	VA	WV					
MD	Baltimore City	NC	OH	PA	VA	WV					
NJ	Union	MD/DC	MI	NY	OH	PA	WV				
NY	New York	MD/DC	OH	PA	WV						
PA	Allegheny	IL	IN	KY	MI	OH	WV				
PA	Beaver	IN	MI	OH	WV						
PA	Berks	MD/DC	MI	NY	OH	VA	WV				
PA	Cambria	IN	MD/DC	MI	OH	WV					
PA	Dauphin	MD/DC	MI	OH	VA	WV					
PA	Delaware	MD/DC	MI	OH	VA	WV					
PA	Lancaster	IN	MD/DC	MI	NY	OH	VA	WV			
PA	Philadelphia	MD/DC	MI	OH	VA	WV					
PA	Washington	IN	KY	MI	OH	WV					
PA	Westmoreland	IN	KY	MD/DC	MI	OH	WV				
PA	York	MD/DC	MI	OH	VA	WV					

Table 2-2. Maximum downwind $PM_{2.5}$ contribution ($\mu g/m^3$) for each of the 37 upwind states (2001 data).

Upwind State	Maximum Downwind Contribution	Upwind State	Maximum Downwind Contribution
Alabama	0.98	Nebraska	0.07
Arkansas	0.19	New Hampshire	<0.05
Connecticut	<0.05	New Jersey	0.13
Delaware	0.14	New York	0.34
Florida	0.45	North Carolina	0.31
Georgia	1.27	North Dakota	0.11
Illinois	1.02	Ohio	1.67
Indiana	0.91	Oklahoma	0.12
Iowa	0.28	Pennsylvania	0.89
Kansas	0.11	Rhode Island	<0.05
Kentucky	0.9	South Carolina	0.4
Louisiana	0.25	South Dakota	<0.05
Maine	<0.05	Tennessee	0.65
Maryland/DC	0.69	Texas	0.29
Massachusetts	0.07	Vermont	<0.05
Michigan	0.62	Virginia	0.44
Minnesota	0.21	West Virginia	0.84
Mississippi	0.23	Wisconsin	0.56
Missouri	1.07		

2.5. Seasonal differences

Eastern and western coastal regions of the United States and Canada show marked seasonality in the concentration and composition of fine particle pollution, while central interior regions do not (NARSTO, 2003). While MANE-VU extends inland as far as the Pennsylvania and Ohio border, the majority of $PM_{2.5}$ NAAQS nonattainment areas and Class I areas affected by the Regional Haze Rule cluster along the East Coast and thus typically show strong seasonal influences. Maximum $PM_{2.5}$ concentrations typically occur during the summer over most of the rural Northeast, with observed summer values for rural areas in the region, on average, twice those of winter. In urban locations, summertime and wintertime $PM_{2.5}$ levels are more comparable and whether one season dominates over the other is more of a function of inter-annual variability of meteorology and fire activity (i.e., summertime fire activity can push average $PM_{2.5}$ values higher in some years). As described below, the reason for the wintertime strength of $PM_{2.5}$ levels in urban areas is related to the greater concentration of local pollution that accumulates when temperature inversions are present, significantly boosting the wintertime $PM_{2.5}$ levels. Winter nitrate concentrations are generally higher than those observed in summer and, as mentioned above, urban concentrations typically exceed rural concentrations year-round. In addition, local mobile source carbon grows in importance during wintertime. Hence, in some large urban areas such as Philadelphia and New York City, peak concentrations of $PM_{2.5}$ can occur in winter.

The conceptual descriptions that explain elevated regional $PM_{2.5}$ peak concentrations in the summer differs significantly from those that explain the largely urban peaks observed during winter. On average, summertime concentrations of sulfate in the northeastern United States are more than twice that of the next most important fine particle constituent, OC, and more than four times the combined concentration of nitrate and black carbon (BC) constituents (NARSTO, 2003). Episodes of high summertime sulfate concentrations are consistent with stagnant meteorological flow conditions upwind of MANE-VU and the accumulation of airborne sulfate (via atmospheric oxidation of SO_2) followed by long-range transport of sulfur emissions from industrialized areas within and outside the region.

National assessments (NARSTO, 2003) have indicated that in the winter, sulfate levels in urban areas are almost twice as high as background sulfate levels across the eastern U.S., indicating that the local urban contribution to wintertime sulfate levels is comparable in magnitude to the regional sulfate contribution from long-range transport. MANE-VU's network analysis for the winter of 2002 suggests that the local enhancement of sulfate in urban areas of MANE-VU is somewhat less with ranges from 25 to 40% and that the long-range transport component of $PM_{2.5}$ sulfate is still the dominant contributor in most eastern cities.

In the winter, urban OC and sulfate each account for about a third of the overall $PM_{2.5}$ mass concentration observed in Philadelphia and New York City. Nitrate also makes a significant contribution to urban $PM_{2.5}$ levels observed in the northeastern United States during the winter months. Wintertime concentrations of OC and NO_3 in urban areas can be twice the average regional concentrations of these pollutants,

indicating the importance of local source contributions (NARSTO, 2003). This is likely because winter conditions are more conducive to the formation of local inversion layers that prevent vertical mixing. Under these conditions, emissions from tailpipe, industrial, and other local sources become concentrated near the Earth's surface, adding to background pollution levels associated with regionally transported emissions.

It is worth noting that while sulfate plays a significant role in episodes of elevated particle pollution during summer and winter months, the processes by which sulfate forms may vary seasonally. Nearly every source apportionment study reviewed by USEPA (2003) identified secondary sulfate originating from coal combustion sources as the largest or one of the largest contributors to overall fine particle mass in the region. It often accounted for more than 50 percent of $PM_{2.5}$ mass at some locations during some seasons. In a few cases, source apportionment studies identified a known local source of sulfate, but most assessments (in conjunction with back trajectory analysis) have pointed to coal-fired power plants in the Midwest as an important source for regional sulfate. Studies with multiple years of data have also tended to identify a distinguishable chemical "signature" for winter versus summer sources of sulfate, with the summer version typically accounting for a greater share of overall fine particle mass. Researchers have speculated that the two profiles represent two extremes in the chemical transformation processes that occur in the atmosphere between the source regions where emissions are released and downwind receptor sites. We note that while coal combustion is often referred to as the "sulfate source" because of the dominance of its sulfate contribution, coal combustion is often a source of significant amounts of organic carbon and is usually the single largest source of selenium (Se) and other heavy metal trace elements (USEPA, 2003).

Figure 2-9. Moving 60-day average of fine aerosol mass concentrations based on long-term data from two northeastern cities

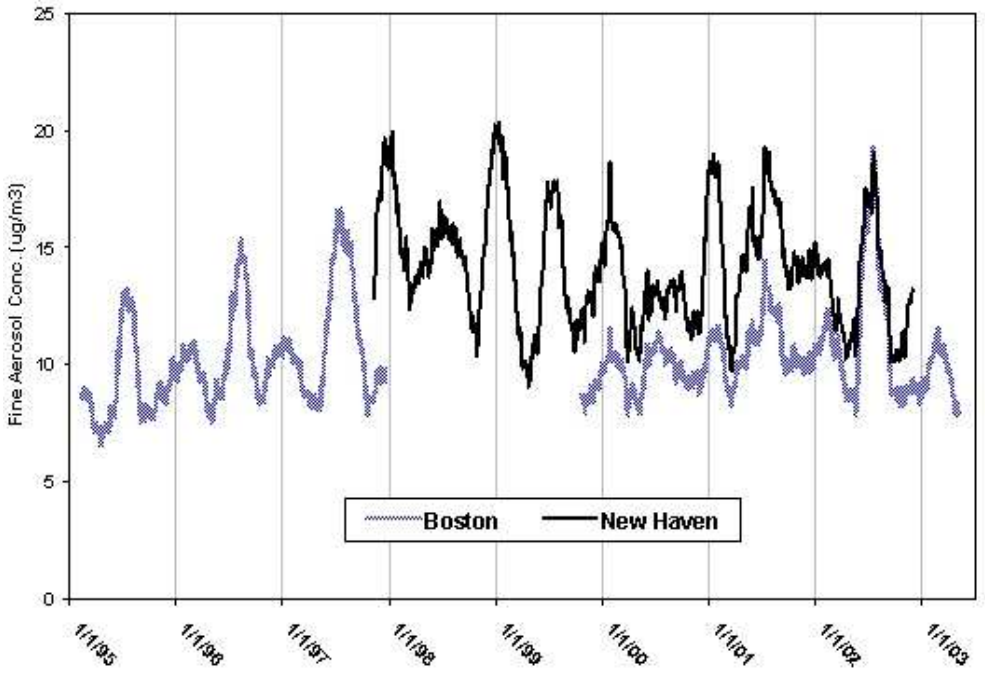
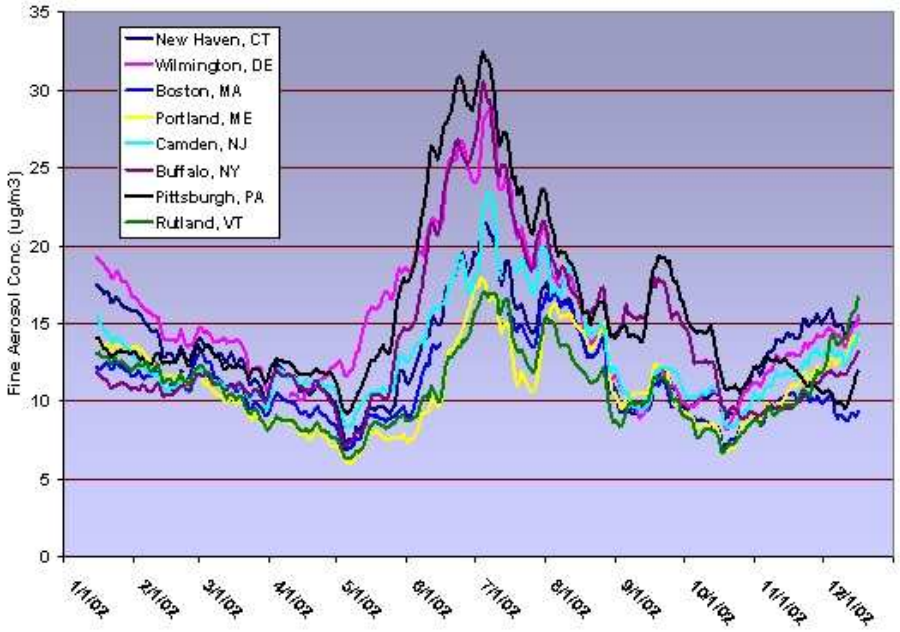


Figure 2-10. The 30-day average $PM_{2.5}$ concentrations from 8 northeastern cities during 2002



In general, fine particle concentrations in MANE-VU are highest during the warmest (summer) months but also exhibit a secondary peak during the coldest (winter) months that can dominate during some years, particularly in urban locations. This bimodal seasonal distribution of peak values is readily apparent in Figure 2-9. The figure shows the smoothed 60-day running average of fine particle mass concentrations using continuous monitoring data from two northeastern cities over a period of several years.

Figure 2-10 also demonstrates this bimodal pattern. Though slightly more difficult to discern in just a single year's worth of data, a "W" pattern does emerge at almost all sites across the region during 2002 with the winter peak somewhat lower than the summer peak at most sites. Urban monitors in Wilmington, Delaware and New Haven, Connecticut have wintertime peak values approaching those of summer.

In the summertime, MANE-VU sites repeatedly experience sulfate events due to transport from regions to the south and west. During such events, both rural and urban sites throughout MANE-VU record high (i.e., >15 $\mu g/m^3$) daily average $PM_{2.5}$ concentrations. Meteorological conditions during the summer frequently allow for summer "stagnation" events when very low wind speeds and warm temperatures (upwind and over MANE-VU) allow pollution levels to build in an air mass as it slowly moves across the continent. During these events, atmospheric ventilation is poor and local emission sources add to the burden of transported pollution with the result that concentrations throughout the region (both rural and urban) are relatively uniform. Generally, there are enough of these events to drive the difference between urban and rural sites down to less than 1 $\mu g/m^3$ during the warm or hot months of the year. As a result, concentrations of fine particles aloft will often be higher than at ground-level during the summertime, especially at rural monitoring sites. Thus, when atmospheric "mixing" occurs during summer[13] mornings (primarily 7 to 11 a.m.), fine particle concentrations at ground-level can actually increase (see Hartford, CT or Camden, NJ in Figure 2-11).

[13] Here we define summer as May, June, July and August.

Figure 2-11. Mean hourly fine aerosol concentrations during 2002 summer months

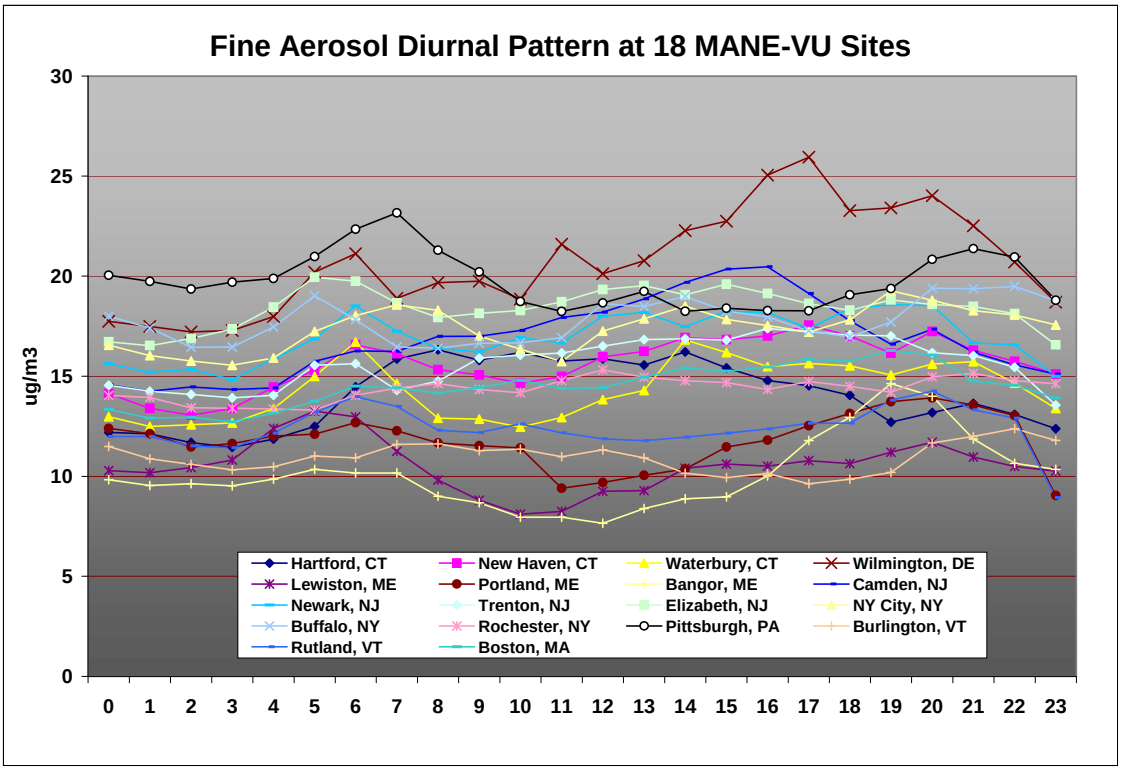
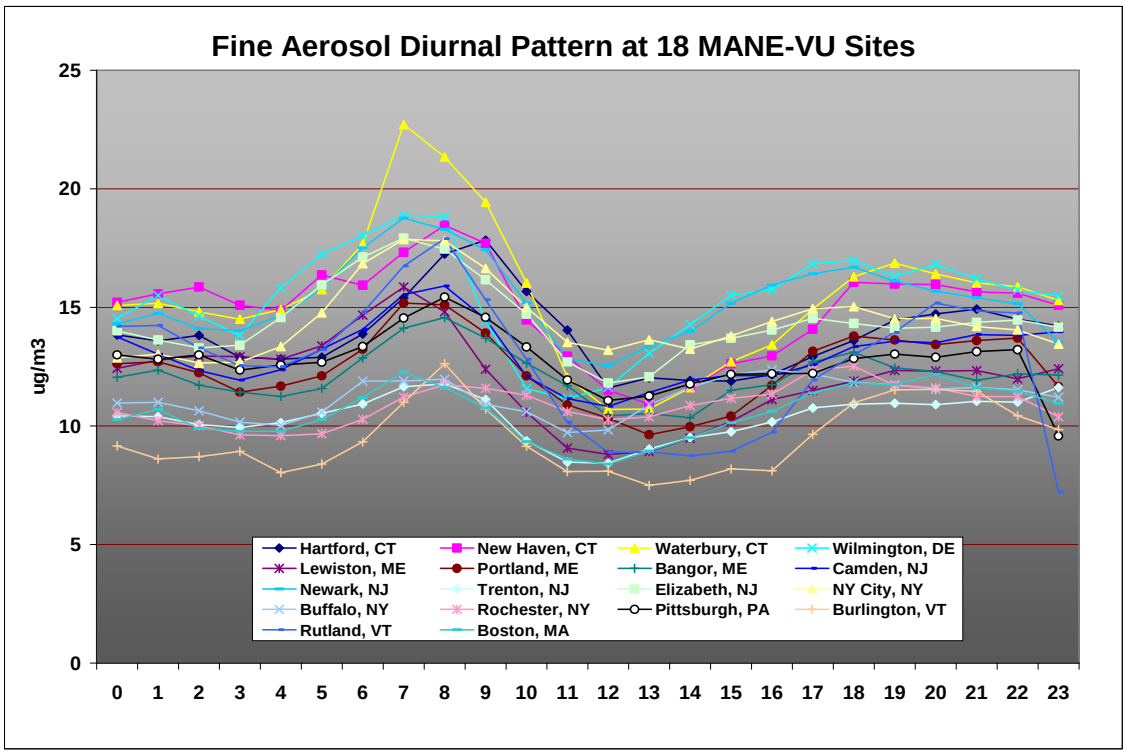


Figure 2-12. Mean hourly fine aerosol concentrations during 2002 winter months



During the wintertime, strong inversions frequently trap local emissions overnight and during the early morning, resulting in elevated urban concentrations. These inversions occur when the Earth's surface loses thermal energy by radiating it into the atmosphere (especially on clear nights). The result is a cold, stable layer of air near the ground. At sunrise, local emissions (both mobile and stationary) begin increasing in strength and build-up in the stable ground layer (which may extend only 100 meters or less above the ground). Increasing solar radiation during the period between 10 a.m. and noon typically breaks this cycle by warming the ground layer so that it can rise and mix with air aloft. Because the air aloft during wintertime is typically less polluted than the surface layer, this mixing tends to reduce ground-level particle concentrations (see Figure 2-12). This diurnal cycle generally drives wintertime particle concentrations, although the occasional persistent temperature inversion can have the effect of trapping and concentrating local emissions over a period of several days, thereby producing a significant wintertime pollution episode.

Rural areas experience the same temperature inversions but have relatively fewer local emissions sources so that wintertime concentrations in rural locations tend to be lower than those in nearby urban areas. Medium and long-range fine particle transport events do occur during the winter but to a far lesser extent than in the summertime. In sum, it is the interplay between local and distant sources together with seasonal meteorological conditions that drives the observed 3–4 $\mu g/m^3$ wintertime urban-rural difference in $PM_{2.5}$ concentrations.

Visually hazy summer days in the Northeast can appear quite different from hazy winter days. The milky, uniform visibility impairment shown in Figure 2-13 is typical of summertime regional haze events in the Northeast. During the winter, by comparison, reduced convection and the frequent occurrence of shallow inversion layers often creates a layered haze with a brownish tinge, as shown in Figure 2-14. This visual difference suggests seasonal variation in the relative contribution of different gaseous and particle constituents during the summer versus winter months (NESCAUM, 2001). Rural and inland areas tend not to experience these layered haze episodes as frequently due to the lack of local emission sources in most rural areas (valleys with high wood smoke contributions are an exception).

Overall (regional) differences in summer versus winter particle mass concentrations and corresponding visibility impairment (as measured by light extinction) are largely driven by seasonal variation in sulfate mass concentrations. This is because winter meteorological conditions are less conducive to the oxidation of sulfate from SO_2 (as borne out by the previously cited source apportionment studies). In addition, seasonal differences in long-range transport patterns from upwind SO_2 source regions may be a factor.

The greater presence of nitrate during the cold season is a consequence of the chemical properties of ammonium nitrate. Ammonia bonds more weakly to nitrate than it does to sulfate, and ammonium nitrate tends to dissociate at higher temperatures. Consequently, ammonium nitrate becomes more stable at lower temperatures and hence contributes more to $PM_{2.5}$ mass and light extinction during the winter months relative to the summer (NESCAUM, 2001).

Figure 2-13. Summertime at Mt. Washington

Clean Day **Typical Haze Event**

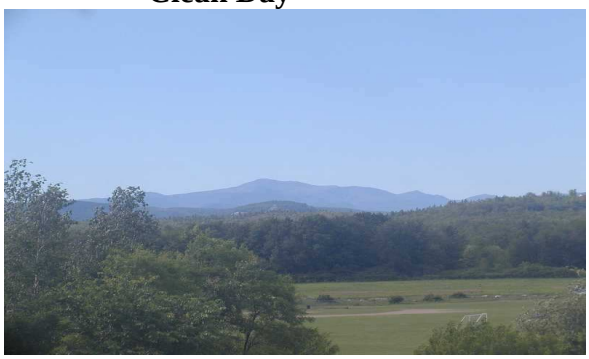
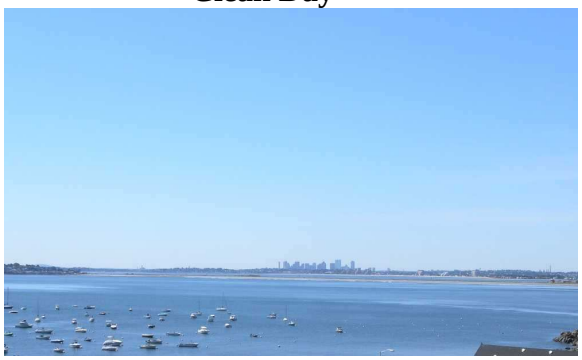


Figure 2-14. Wintertime in Boston

Clean Day **Typical Haze Event**



2.6. Summary

The presence of fine particulate matter in ambient air significantly degrades public health and obscures visibility during most parts of the year at sites across the MANE-VU region. Particle pollution generally, and its sulfate component specifically, constitute the principle driver for regional visibility impacts. While the broad region experiences visibility impairment, it is most severe in the southern and western portions of MANE-VU that are closest to large power plant SO_2 sources in the Ohio River and Tennessee Valleys.

Summer visibility impairment is driven by the presence of regional sulfate, whereas winter visibility depends on a combination of regional and local influences coupled with local meteorological conditions (inversions) that lead to the concentrated build-up of pollution.

Sulfate is the key particle constituent from the standpoint of designing control strategies to improve visibility conditions in the northeastern United States. Significant further reductions in ambient sulfate levels are achievable, though they will require more than proportional reductions in SO_2 emissions.

Long-range pollutant transport and local pollutant emissions are important, especially along the eastern seaboard, so one must also look beyond the achievement of further sulfate reductions. During the winter months, in particular, consideration also needs to be given to reducing urban sources of SO_2, NO_X and OC (NARSTO, 2003) .

References

Husar, R.B. and W.E. Wilson, "Haze and Sulfur Emission Trends in the Eastern United States", *Environ. Sci. Technol.* 1993, *27* (1), 12-16.

Malm, W.C., B.A. Schichtel, R. B. Ames, and K. A. Gebhart (2002), "A 10-year spatial and temporal trend of sulfate across the United States," *J. Geophys. Res*. 107(D22):4627, doi:10.1029/2002JD002107.

NARSTO, *Particulate Matter Science for Policy Makers: A NARSTO Assessment*, EPRI 1007735, February, 2003.

NESCAUM, *Regional Haze and Visibility Impairment in the Northeast and Mid-Atlantic United States*, Northeast States for Coordinated Air Use Management, Boston, MA, January, 2001.

NESCAUM, *2002: A Year in Review*, Northeast States for Coordinated Air Use Management, Boston, MA, December, 2004b.

NESCAUM, *Contributions to Regional Haze and Visibility Impairment in the Northeast and Mid-Atlantic United States*, Northeast States for Coordinated Air Use Management, Boston, MA, June, 2006.

USEPA, *COMPILATION OF EXISTING STUDIES ON SOURCE APPORTIONMENT FOR $PM_{2.5}$, Second Draft*, Emissions, Monitoring, and Analysis Division, Office of Air Quality Planning and Standards, Research Triangle Park, NC, August, 2003.

USNPS, *Assessment of Air Quality and Related Values in Shenandoah National Park*, Technical Report NPS/NERCHAL/NRTR-03/090, U.S. Department of the Interior, National Park Service, Northeast Region, Philadelphia, Pennsylvania, May, 2003.

Watson, J., "Visibility: Science and Regulation", *JAWMA* 52:628-713, 2002.

West, J.J., Ansari, A.S., and S. N. Pandis, "Marginal $PM_{2.5}$: Nonlinear Aerosol Mass Response to Sulfate Reductions in the Eastern United States." *JAWMA* 49:1415-1424, 1999.

USEPA, "Technical Support Document for the Final Clean Air Interstate Rule, Air Quality Modeling." U.S. Environmental Protection Agency Office of Air Quality Planning and Standards. March 2005

.

3. MANE-VU EMISSION INVENTORY CHARACTERISTICS FOR FINE PARTICLES

The pollutants that affect fine particle formation and visibility are sulfur oxides (SO_X), NO_X, VOCs, ammonia (NH_3), and particles with an aerodynamic diameter less than or equal to 10 and 2.5 µm (i.e., primary PM_{10} and $PM_{2.5}$). The emissions dataset illustrated in this section is the 2002 MANE-VU Version 2 regional haze emissions inventory. The MANE-VU regional haze emissions inventory version 3.0, released in April 2006, has superseded version 2 for modeling purposes.

3.1. Emissions inventory characteristics

3.1.1. Sulfur dioxide (SO_2)

SO_2 is the primary precursor pollutant for sulfate particles. Ammonium sulfate particles are the largest contributor to $PM_{2.5}$ mass on an annual average basis at MANE-VU nonattainment sites. It also accounts for more than 50 percent of particle-related light extinction at northeastern Class I areas on the clearest days and for as much as or more than 80 percent on the haziest days. Hence, SO_2 emissions are an obvious target of opportunity for both addressing $PM_{2.5}$ nonattainment and for reducing regional haze in the eastern United States. Combustion of coal and, to a substantially lesser extent, of certain petroleum products accounts for most anthropogenic SO_2 emissions. In fact, in 1998 a single source category — coal-burning power plants — was responsible for two-thirds of total SO_2 emissions nationwide (NESCAUM, 2001).

Figure 3-1 shows SO_2 emissions trends in MANE-VU states[14] extracted from the National Emissions Inventories (NEI) for the years 1996, 1999 (MARAMA, 2004), and the 2002 MANE-VU inventory. Most of the states (with the exception of Maryland) show declines in year 2002 annual SO_2 emissions as compared to 1996 emissions. Some of the states show an increase in 1999 followed by a decline in 2002 and others show consistent declines throughout the entire period. The upward trend in emissions after 1996 probably reflects electricity demand growth during the late 1990s combined with the availability of banked SO_2 emissions allowances from initial over-compliance with control requirements in Phase 1 of the USEPA Acid Rain Program. This led to relatively low market prices for allowances later in the decade, which encouraged utilities to purchase allowances rather than implement new controls as electricity output expanded. The observed decline in the 2002 SO_2 emissions inventory reflects implementation of the second phase of the USEPA Acid Rain Program, which in 2000 further reduced allowable emissions and extended emissions limits to more power plants.

Figure 3-2 shows the percent contribution from different source categories to overall annual 2002 SO_2 emissions in MANE-VU states. The chart shows that point sources dominate SO_2 emissions, which primarily consist of stationary combustion sources for generating electricity, industrial energy, and heat. Smaller stationary combustion sources called “area sources” (primarily commercial and residential heating)

[14] The description of MANE-VU state inventories discussed throughout this section does not include the portion of Virginia in the Washington, DC metropolitan area.

are another important source category in MANE-VU states. By contrast, on-road and non-road mobile sources make only a relatively small contribution to overall SO_2 emissions in the region (NESCAUM, 2001).

Figure 3-1. State level sulfur dioxide emissions

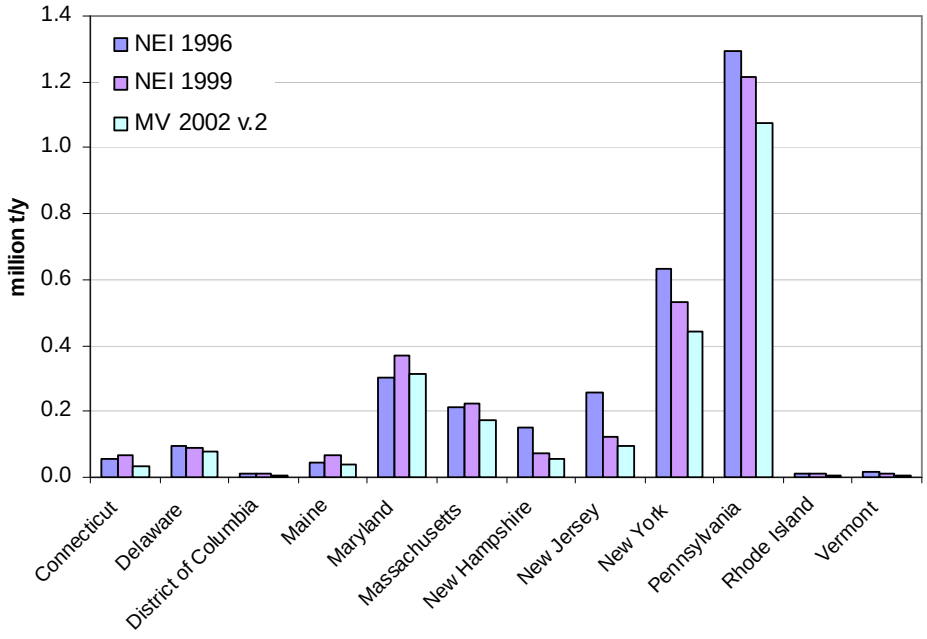


Figure 3-2. 2002 MANE-VU state SO_2 inventories

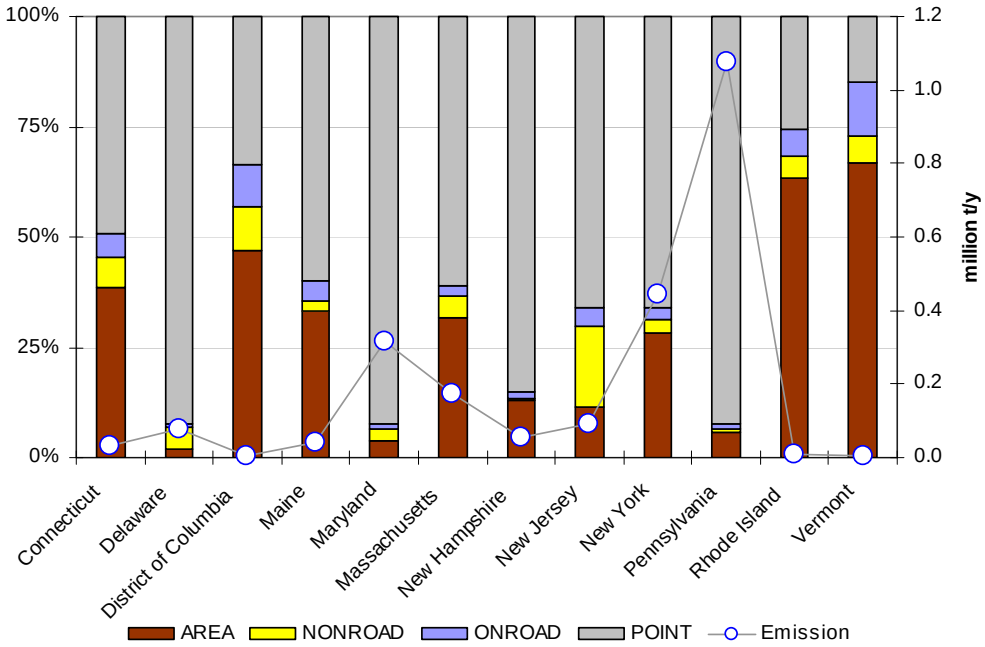


Figure Key: Bars = Percentage fractions of four source categories; Circles = Annual emissions amount in 10^6 tons per year. Note that Version 2 of the MANE-VU inventory was used and the Virginia portion of the Washington, DC metropolitan area is not shown in the figure.

3.1.2. Volatile organic compounds (VOCs)

Existing emission inventories generally refer to VOCs based on their historical contribution to ozone formation. From a fine particle perspective, VOCs (also referred to as hydrocarbons) are of concern because they can react in the atmosphere to form secondary organic aerosol (SOA) as a result of condensation and oxidation processes. The SOA component of fine particles also obscures visibility, but this component has a smaller impact on visibility (on a per unit mass basis) relative to sulfate or nitrate, which have an affinity for water that allows them to significantly "grow" as particles under humid conditions. Nonetheless, organic carbon typically has the second largest visibility impact at most Class I sites next to sulfate, given its large mass contribution.

As shown in Figure 3-3, the VOC inventory is dominated by mobile and area sources. Most VOC emissions in MANE-VU, however, come from natural sources, which are not shown in the figure. Among the human-caused VOC emissions, on-road mobile sources of VOCs include exhaust emissions from gasoline passenger vehicles and diesel-powered heavy-duty vehicles as well as evaporative emissions from transportation fuels. VOC emissions may also originate from a variety of area sources (including solvents, architectural coatings, and dry cleaners) as well as from some point sources (e.g., industrial facilities and petroleum refineries).

Naturally occurring (biogenic) VOC emissions are caused by the release of natural organic compounds from plants in warm weather. Natural, or biogenic, VOCs contribute significantly to fine particle formation. Biogenic VOCs are not included in Figure 3-3, but nationally, they represent roughly two-thirds of all annual VOC emissions (USEPA, 2006). Biogenic emissions are extremely difficult to estimate, as it requires modeling the behavior of many plants as well as their responses to the environment.

With regard to fine particle formation, understanding the transport dynamics and source regions for organic carbon is likely to be more complex than for sulfate. This is partly because of the large number and variety of VOC species, the fact that their transport characteristics vary widely, and the fact that a given species may undergo numerous complex chemical reactions in the atmosphere. Thus, the organic carbon contribution to fine particles in the East is likely to include manmade pollution transported from a distance, manmade pollution from nearby sources, and biogenic emissions, especially terpenes from coniferous forests.

For fine particles derived from organic carbon, the oxidation of hydrocarbon molecules containing seven or more carbon atoms is generally the most significant pathway for their formation (Odum *et al.*, 1997). Recent research, however, suggests that smaller reactive hydrocarbons like isoprene not only contribute significantly to ground-level ozone, which may indirectly impact organic aerosol formation, but also contribute directly to ambient organic aerosol through heterogeneous processes (Claeys et al., 2004; Kroll et al., 2005).

Figure 3-3. 2002 MANE-VU state VOC inventories

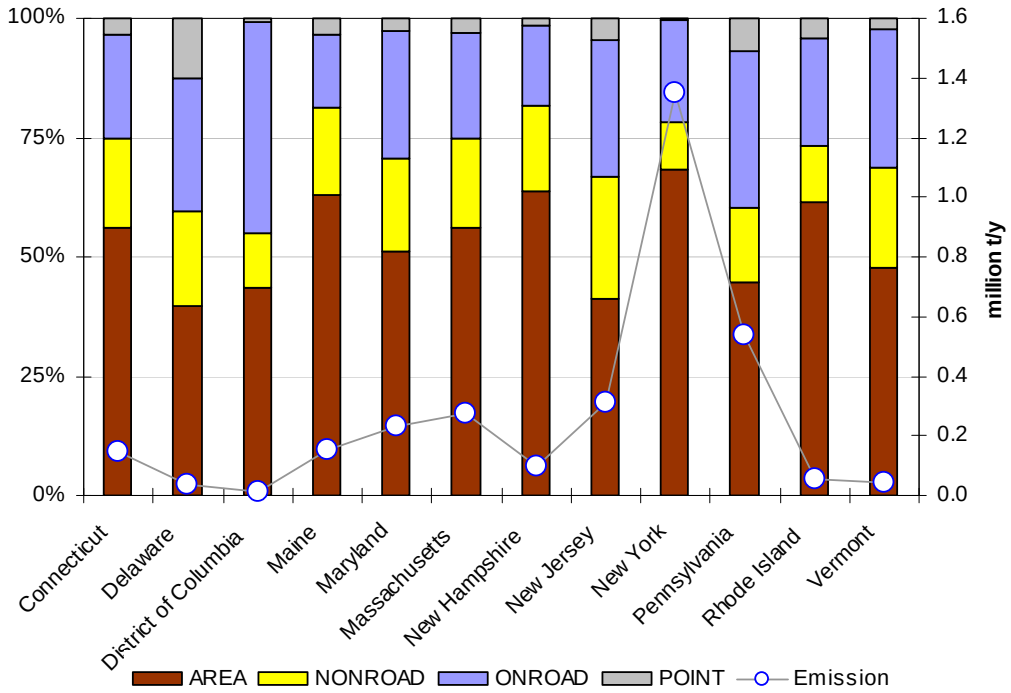


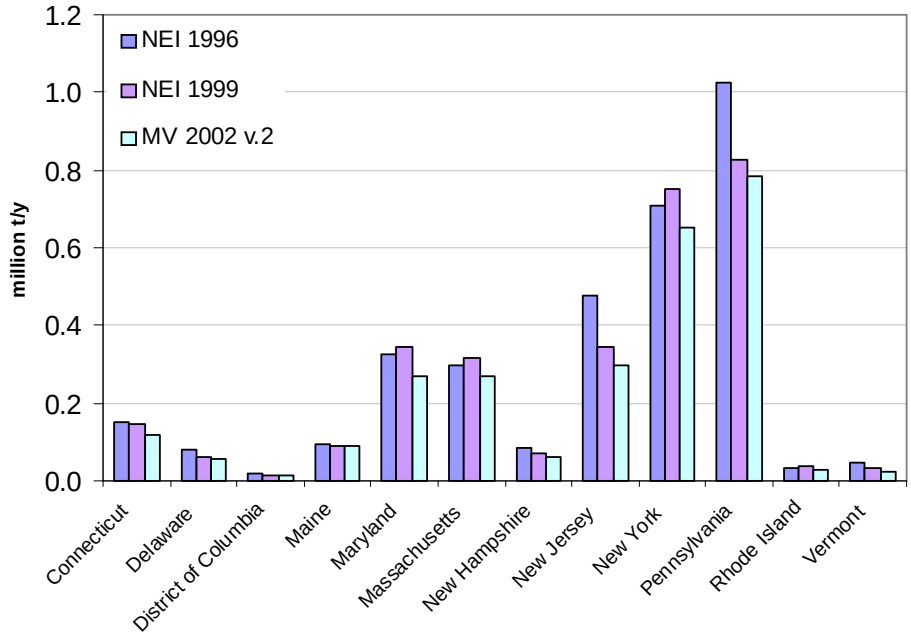
Figure key: Bars = Percentage fractions of four source categories; Circles = Annual emissions amount in 10^6 tons per year. Note that Version 2 of the MANE-VU inventory was used and the Virginia portion of the Washington, DC metropolitan area is not shown in the figure. Biogenic VOCs are not included in this figure.

3.1.3. Oxides of nitrogen (NO_X)

NO_X emissions contribute directly to $PM_{2.5}$ nonattainment and visibility impairment in the eastern U.S. by forming nitrate particles. Nitrate generally accounts for a substantially smaller fraction of fine particle mass and related light extinction than sulfate and organic carbon regionally in MANE-VU. Notably, nitrate may play a more important role at urban sites and in the wintertime. In addition, NO_X may have an indirect effect on summertime visibility by virtue of its role in the formation of ozone, which in turn promotes the formation of secondary organic aerosols (NESCAUM, 2001).

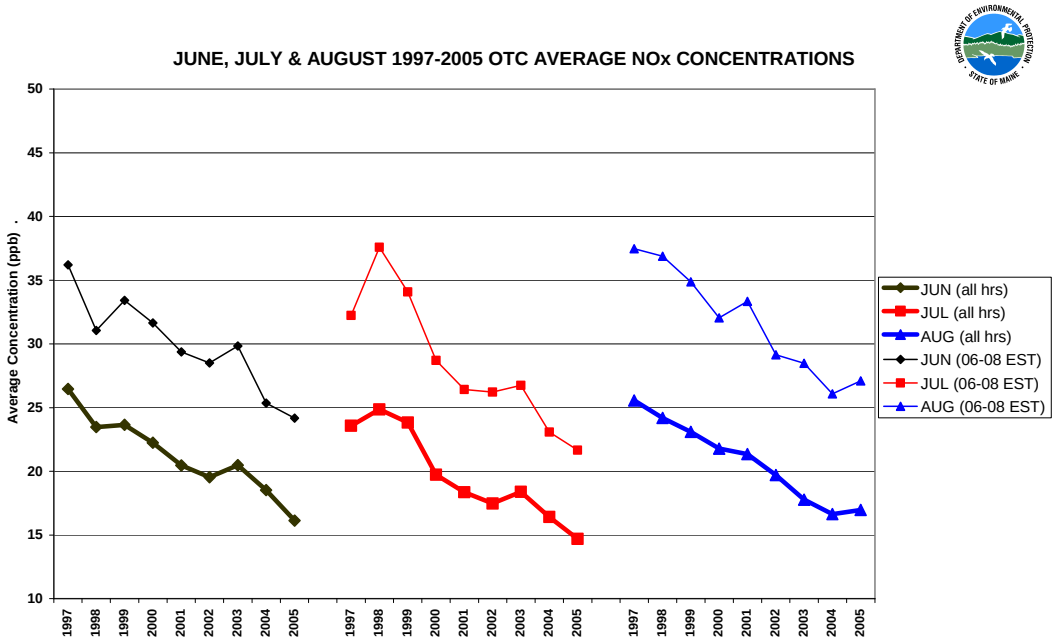
Figure 3-4 shows NO_X emissions in MANE-VU at the state level. Since 1980, nationwide emissions of NO_X from all sources have shown little change. In fact, emissions increased by 2 percent between 1989 and 1998 (USEPA, 2000a). This increase is most likely due to industrial sources and the transportation sector, as power plant combustion sources have implemented modest emissions reductions during the same time period. Most states in MANE-VU experienced declining NO_X emissions from 1996 through 2002, except Massachusetts, Maryland, New York, and Rhode Island, which show an increase in NO_X emissions in 1999 before declining to levels below 1996 emissions in 2002.

Figure 3-4. State level nitrogen oxides emissions



Monitored ambient NO_X trends during the summer from 1997 to 2005 corroborate the downward trend in NO_X emissions seen in the emissions inventories for MANE-VU. As seen in Figure 3-5, the 24-hour (lower trend lines) and 6 a.m.-8 a.m. (upper trend lines) NO_X concentrations indicate decreases in NO_X over this time period in MANE-VU. The NO_X reductions likely come from decreasing vehicle NO_X emissions due to more stringent motor vehicle standards as well as NO_X reductions from MANE-VU NO_X Budget Program and the NO_X SIP Call (mainly power plants).

Figure 3-5. Plot of monitored NO_X trends in MANE-VU during 1997-2005



Note: Upper trend lines correspond to NO_X measured from 0600-0800 EST in the morning. Lower trend lines correspond to NO_X measured over entire day (created by Tom Downs, Maine Department of Environmental Protection).

Power plants and mobile sources generally dominate state and national NO_X emissions inventories. Nationally, power plants account for more than one-quarter of all NO_X emissions, amounting to over six million tons. The electric sector plays an even larger role, however, in parts of the industrial Midwest where high NO_X emissions have a particularly significant power plant contribution. By contrast, mobile sources dominate the NO_X inventories for more urbanized mid-Atlantic and New England states to a far greater extent, as shown in Figure 3-6. In these states, on-road mobile sources — a category that mainly includes highway vehicles — represent the most significant NO_X source category. Emissions from non-road (i.e., off-highway) mobile sources, primarily diesel-fired engines, also represent a substantial fraction of the inventory.

Figure 3-6. 2002 MANE-VU state NO_X inventories

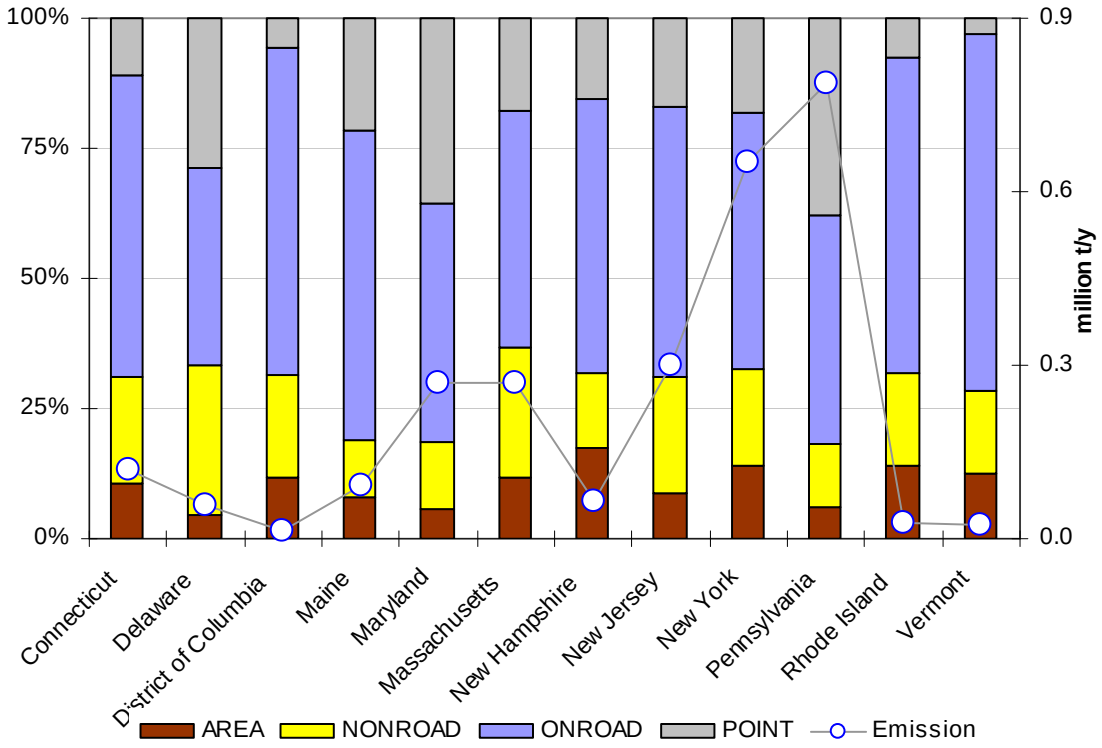


Figure key: Bars = Percentage fractions of four source categories; Circles = Annual emissions amount in 10^6 tons per year. Note that Version 2 of the MANE-VU inventory was used and the Virginia portion of the Washington, DC metropolitan area is not shown in the figure.

3.1.4. Primary particulate matter (PM_{10} and $PM_{2.5}$)

Directly-emitted or "primary" particles (as distinct from secondary particles that form in the atmosphere through chemical reactions involving precursor pollutants like SO_2 and NO_X) also contribute to fine particle levels in the atmosphere. For regulatory purposes, we make a distinction between particles with an aerodynamic diameter less than or equal to 10 micrometers and smaller particles with an aerodynamic diameter less than or equal to 2.5 micrometers (i.e., primary PM_{10} and $PM_{2.5}$, respectively).

Figure 3-7 and Figure 3-8 show PM_{10} and $PM_{2.5}$ emissions for MANE-VU states for the years 1996, 1999, and 2002. Note that, as opposed to the other constituents of PM, the 2002 inventory values for PM_{10} are drawn from the 2002 NEI. Most states show a steady decline in annual PM_{10} emissions over this time period. By contrast, emission trends for primary $PM_{2.5}$ are more variable.

Crustal sources are significant contributors of primary PM emissions. This category includes fugitive dust emissions from construction activities, paved and unpaved roads, and agricultural tilling. Typically, monitors estimate PM_{10} emissions from these types of sources by measuring the horizontal flux of particulate mass at a fixed downwind sampling location within perhaps 10 meters of a road or field. Comparisons between estimated emission rates for fine particles using these types of measurement techniques and observed concentrations of crustal matter in the ambient air at downwind receptor sites suggest that physical or chemical processes remove a significant fraction of crustal material relatively quickly. As a result, it rarely entrains into layers of the atmosphere where it can transport to downwind receptor locations. Because of this discrepancy between estimated emissions and observed ambient concentrations, modelers typically reduce estimates of total $PM_{2.5}$ emissions from all crustal sources by applying a factor of 0.15 to 0.25 before including in modeling analyses.

From a regional haze perspective, crustal material generally does not play a major role. On the 20 percent best-visibility days during the baseline period (2000-2004), it accounted for 6 to 11 percent of particle-related light extinction at MANE-VU Class 1 sites. On the 20 percent worst-visibility days, however, crustal material generally plays a much smaller role relative to other haze-forming pollutants, ranging from 2 to 3 percent. Moreover, the crustal fraction includes material of natural origin (such as soil or sea salt) that is not targeted under USEPA's Regional Haze Rule. Of course, the crustal fraction can be influenced by certain human activities, such as construction, agricultural practices, and road maintenance (including wintertime salting) — thus, to the extent that these types of activities are found to affect visibility at northeastern Class I sites, control measures targeted at crustal material may prove beneficial.

Experience from the western United States, where the crustal component has generally played a more significant role in driving overall particulate levels, may be helpful where it is relevant in the eastern context. In addition, a few areas in the Northeast, such as New Haven, Connecticut and Presque Isle, Maine, have some experience with the control of dust and road-salt as a result of regulatory obligations stemming from their past nonattainment status with respect to the NAAQS for PM_{10}.

Current emissions inventories for the entire MANE-VU area indicate residential wood combustion represents 25 percent of primary fine particulate emissions in the region. This implies that rural sources can play an important role in addition to the contribution from the region's many highly populated urban areas. An important consideration in this regard is that residential wood combustion occurs primarily in the winter months, while managed or prescribed burning activities occur largely in other seasons. The latter category includes agricultural field-burning activities, prescribed burning of forested areas, and other burning activities such as construction waste burning. Limiting burning to times when favorable meteorological conditions can efficiently disperse resulting emissions can manage many of these types of sources.

Figure 3-7. State level primary PM_{10} emissions

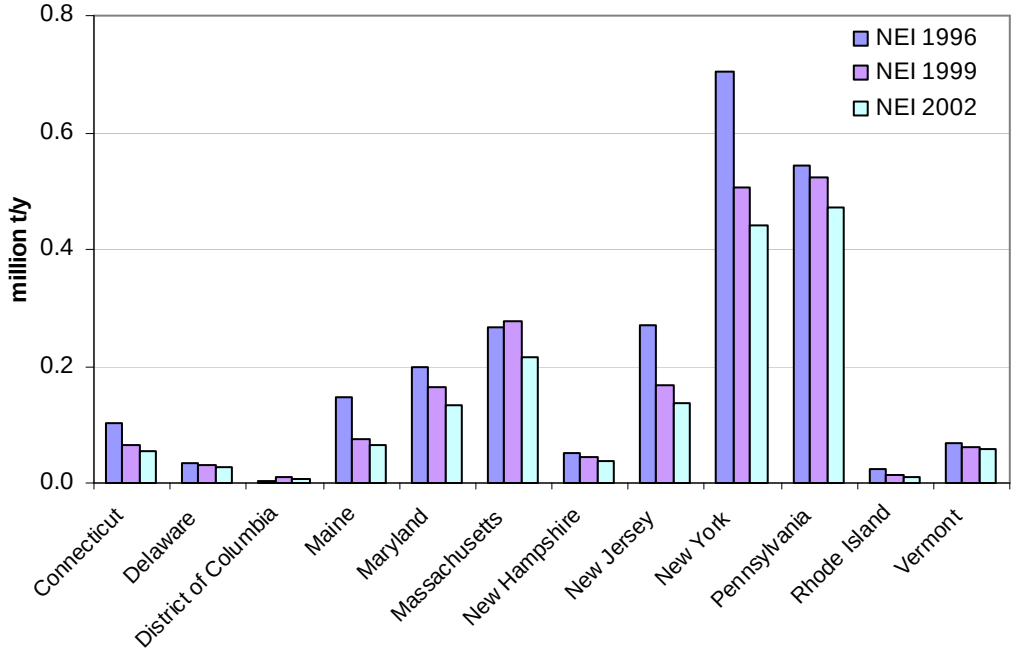


Figure 3-8. State level primary $PM_{2.5}$ emissions

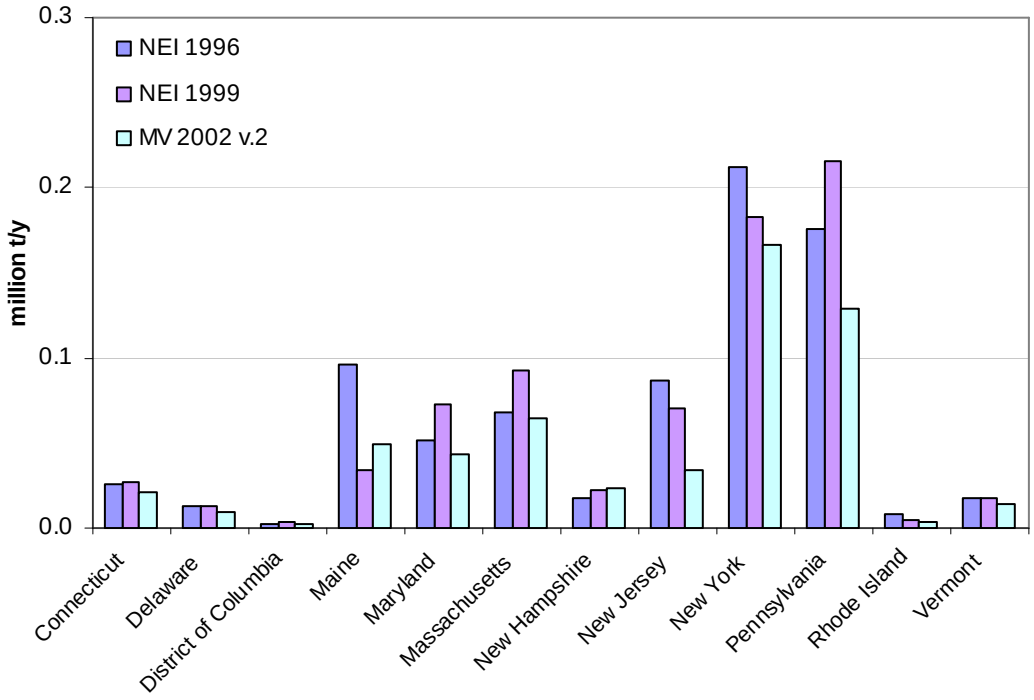


Figure 3-9 and Figure 3-10 show that area and mobile sources dominate primary PM emissions. (The NEI inventory categorizes residential wood combustion and some other combustion sources as area sources.) The relative contribution of point sources is larger in the primary $PM_{2.5}$ inventory than in the primary PM_{10} inventory since the crustal

component (which consists mainly of larger or "coarse-mode" particles) contributes mostly to overall PM_{10} levels. At the same time, pollution control equipment commonly installed at large point sources is usually more efficient at capturing coarse-mode particles.

Figure 3-9. 2002 MANE-VU state primary PM_{10} inventories

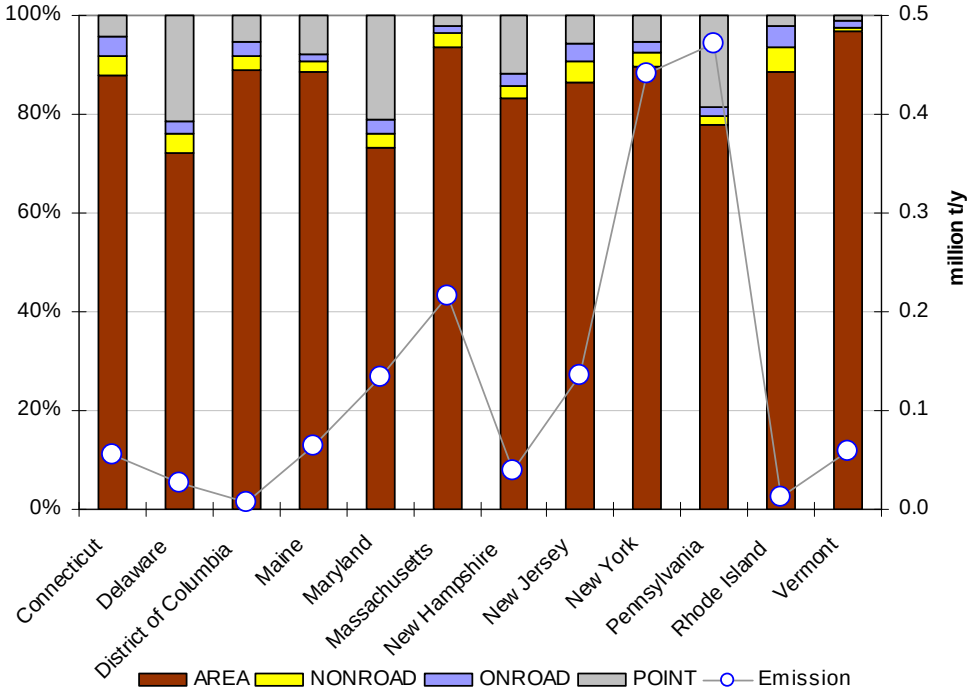


Figure 3-10. 2002 MANE-VU state primary $PM_{2.5}$ inventories

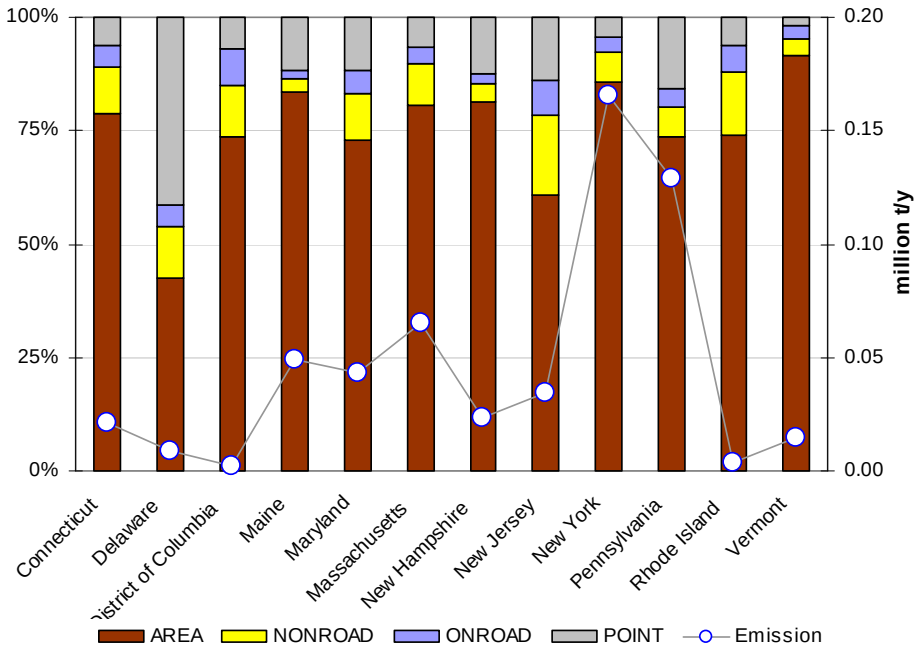


Figure key: Bars = Percentage fractions of four source categories; Circles = Annual emissions amount in 10^6 tons per year. Note that Version 2 of the MANE-VU inventory was used and the Virginia portion of the Washington, DC metropolitan area is not shown in the figure.

3.1.5. Ammonia emissions (NH_3)

Knowledge of ammonia emission sources will be necessary in developing effective regional haze reduction strategies because of the importance of ammonium sulfate and ammonium nitrate in determining overall fine particle mass and light scattering. According to 1998 estimates, livestock and agriculture fertilizer use accounted for approximately 85 percent of all ammonia emissions to the atmosphere (USEPA, 2000b). We need, however, better ammonia inventory data for the photochemical models used to simulate fine particle formation and transport in the eastern United States. Because the USEPA does not regulate ammonia as a criteria pollutant or as a criteria pollutant precursor, these data do not presently exist at the same level of detail or certainty as for NO_X and SO_2.

Ammonium ion (formed from ammonia emissions to the atmosphere) is an important constituent of airborne particulate matter, typically accounting for 10–20 percent of total fine particle mass. Reductions in ammonium ion concentrations can be extremely beneficial because a more-than-proportional reduction in fine particle mass can result. Ansari and Pandis (1998) showed that a one $\mu g/m^3$ reduction in ammonium ion could result in up to a four $\mu g/m^3$ reduction in fine particulate matter. Decision makers, however, must weigh the benefits of ammonia reduction against the significant role it plays in neutralizing acidic aerosol. SO_2 reacts in the atmosphere to form sulfuric acid (H_2SO_4). Ammonia can partially or fully neutralize this strong acid to form ammonium bisulfate or ammonium sulfate. If planners focus future control strategies on ammonia and do not achieve corresponding SO_2 reductions, fine particles formed in the atmosphere will be substantially more acidic than those presently observed.

To address the need for improved ammonia inventories, MARAMA, NESCAUM and USEPA funded researchers at Carnegie Mellon University (CMU) in Pittsburgh to develop a regional ammonia inventory system (Davidson et al., 1999). This study focused on three issues with respect to current emissions estimates: (1) a wide range of ammonia emission factor values, (2) inadequate temporal and spatial resolution of ammonia emissions estimates, and (3) a lack of standardized ammonia source categories.

Figure 3-11 shows that estimated ammonia emissions were fairly stable in the 1996, 1999, and 2002 NEI for MANE-VU states, with some increases observed for Massachusetts, New Jersey and New York. Area and on-road mobile sources dominate the ammonia inventory, according to Figure 3-12. Specifically, emissions from agricultural sources and livestock production account for the largest share of estimated ammonia emissions in MANE-VU, except in the District of Columbia. The two remaining sources with a significant emissions contribution are wastewater treatment systems and gasoline exhaust from highway vehicles.

Figure 3-11. State level ammonia emissions

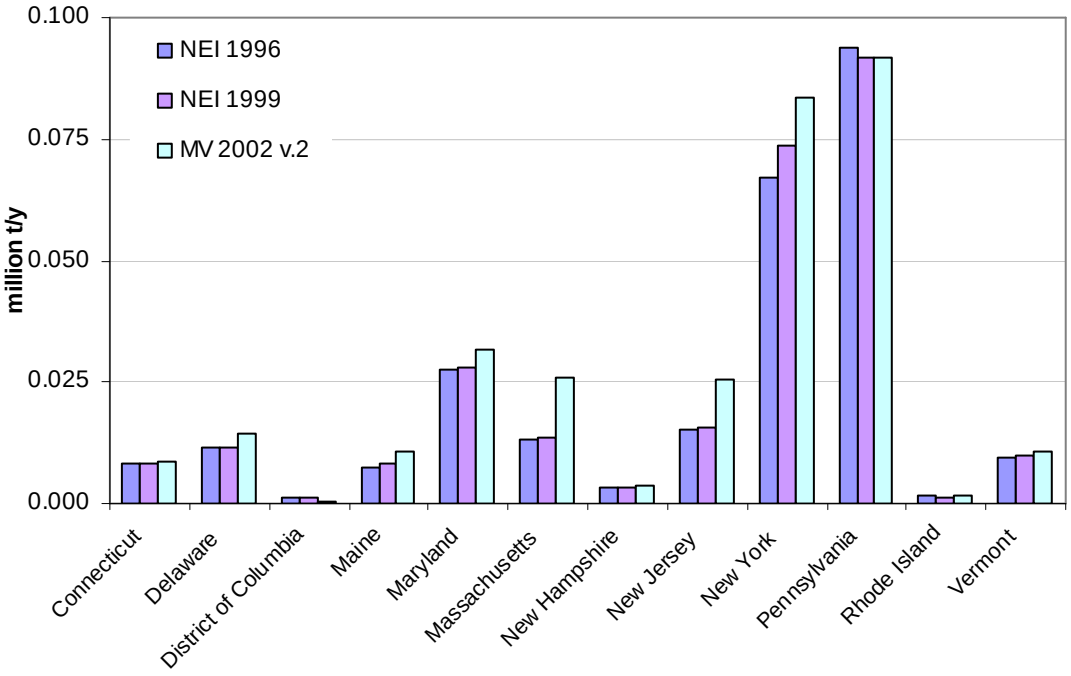


Figure 3-12. 2002 MANE-VU state NH_3 inventories

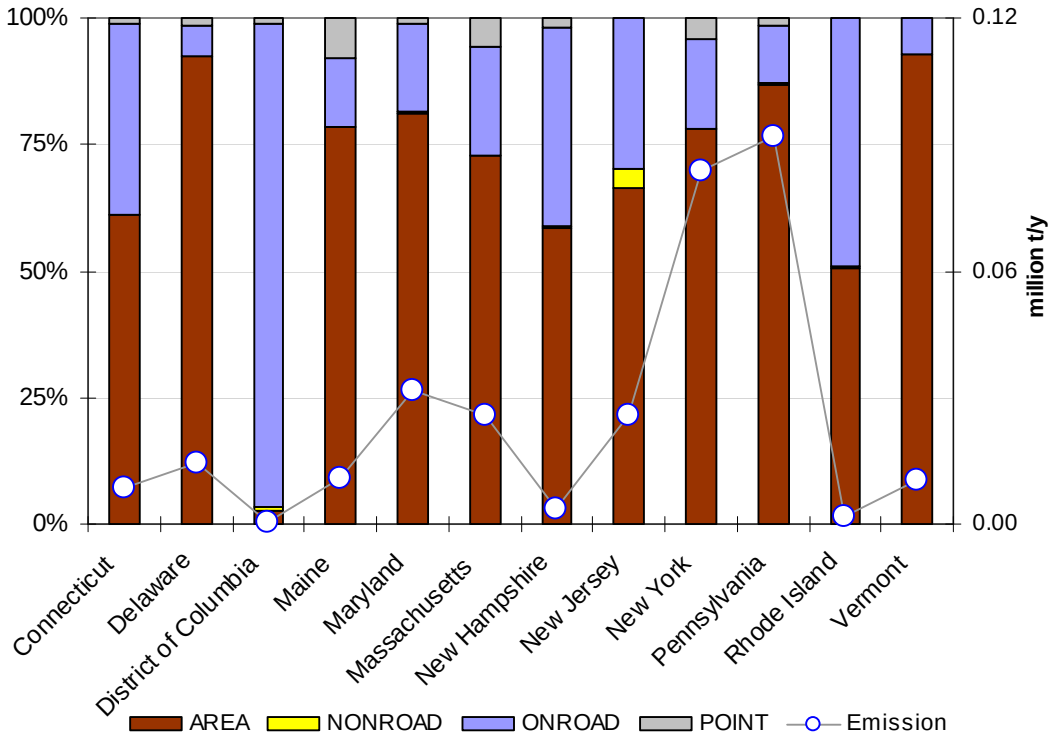


Figure key: Bars = Percentage fractions of four source categories; Circles = Annual emissions amount in 10^6 tons per year. Note that Version 2 of the MANE-VU inventory was used and the Virginia portion of the Washington, DC metropolitan area is not shown in the figure.

3.2. Emissions inventory characteristics outside MANE-VU

SO_2, NO_X and VOC emissions from within MANE-VU are only one component of the emissions contributing to fine particles affecting the MANE-VU region. As regional modeling for the CAIR has shown, emission sources, primarily of SO_2 and NO_X, located outside MANE-VU can significantly contribute to particle sulfate and nitrate transported into the MANE-VU region. Here we present regional emissions information grouped by the three eastern RPOs – MANE-VU, VISTAS (Visibility Improvement State and Tribal Association of the Southeast), and the MWRPO (Midwest RPO). Table 3-1 lists the states in each RPO.

The inventory information is extracted from the USEPA final 2002 National Emissions Inventory (NEI). For consistency, the MANE-VU information here also comes from the 2002 NEI rather than from the MANE-VU Version 2 regional haze emissions inventory described in Section 3.1. The differences between the inventories are not great, as the NEI and the MANE-VU Version 2 inventory are both based on the same inventory information provided by the states.

Table 3-1. Eastern U.S. RPOs and their state members

RPO	State
MWRPO	Illinois
MWRPO	Indiana
MWRPO	Michigan
MWRPO	Ohio
MWRPO	Wisconsin
MANE-VU	Connecticut
MANE-VU	Delaware
MANE-VU	District of Columbia
MANE-VU	Maine
MANE-VU	Maryland
MANE-VU	Massachusetts
MANE-VU	New Hampshire
MANE-VU	New Jersey
MANE-VU	New York
MANE-VU	Pennsylvania
MANE-VU	Rhode Island
MANE-VU	Vermont
VISTAS	Alabama
VISTAS	Florida
VISTAS	Georgia
VISTAS	Kentucky
VISTAS	Mississippi
VISTAS	North Carolina
VISTAS	South Carolina
VISTAS	Tennessee
VISTAS	Virginia
VISTAS	West Virginia

Table 3-2 presents SO_2 emissions by source sector and RPO for the eastern United States. The NO_X emissions by source sector and RPO are presented in Table 3-3 and VOC emissions in Table 3-4. Regionally, SO_2 emissions are more important with respect to regional particle formation and transport. NO_X emissions play an important role in determining the equilibrium between ammonium sulfate and ammonium nitrate formation, especially during winter. VOC emissions contribute to secondary organic aerosol formation.

Table 3-2. SO_2 emissions in eastern RPOs (tons/yr)

RPO	Point	Area	On-road	Non-road	Total
MWRPO	3,336,967	133,415	49,191	82,307	3,601,880
MANE-VU	1,924,573	353,176	39,368	74,566	2,391,683
VISTAS	4,349,437	448,023	83,001	91,307	4,971,769

Table 3-3. NO_X emissions in eastern RPOs (tons/yr)

RPO	Point	Area	On-road	Non-road	Total
MWRPO	1,437,284	184,790	1,290,178	723,844	3,636,096
MANE-VU	680,975	268,997	1,297,357	534,454	2,781,783
VISTAS	2,094,228	266,848	2,160,601	812,615	5,334,293

Table 3-4. VOC emissions in eastern RPOs (tons/yr)

RPO	Point	Area	On-road	Non-road	Total
MWRPO	234,938	1,182,186	660,010	492,027	2,569,160
MANE-VU	93,691	1,798,158	793,541	494,115	3,179,504
VISTAS	458,740	2,047,359	1,314,979	609,539	4,430,617

References

Ansari, A.S., and S.N. Pandis. "Response of inorganic PM to precursor concentrations." *Environ. Sci. Technol*, 32, 2706-2714, 1998.

Claeys, M., W. Wang, A.C. Ion, I. Kourtchev, A. Gelencser, and W. Maenhaut, "Formation of secondary organic aerosols from isoprene and gas-phase oxidation products through reaction with hydrogen peroxide." *Atmos. Environ.*, 38, 4093-4098, 2004.

Davidson, C., R. Strader, S. Pandis, and A. Robinson. *Preliminary Proposal to MARAMA and NESCAUM: Development of an Ammonia Emissions Inventory for the Mid-Atlantic States and New England*. Carnegie Mellon University, Pittsburgh, PA, January 7, 1999.

Kroll, J.H., N.L. Ng, S.M. Murphy, R.C. Flagan, and J.H. Seinfeld. "Secondary Organic Aerosol Formation from Isoprene Photooxidation." *Environ. Sci. Technol.* **40**, 1869-1877, 2006.

MARAMA 2004, http://www.marama.org/visibility/2002%20NEI/index.html

MARAMA. "A Guide to Mid-Atlantic Regional Air Quality," October 2005.

NESCAUM. *Regional Haze and Visibility in the Northeast and Mid-Atlantic States.* NESCAUM, Boston, MA, January 2001.

Odum, J.R., T.P.W. Jungkamp, R.J. Griffin, R.C. Flagan, and J.J. Seinfeld. "The atmospheric aerosol-forming potential of whole gasoline vapor." *Science* **276**, 96-99, 1997.

USEPA. *National Air Quality and Emission Trends Report, 1998*, EPA 454/R-00-003, available online: http://www.epa.gov/oar/aqtrnd98/, 2000a.

USEPA. *National Air Pollutant Trends, 1900 – 1998*, EPA 454/R-00-002, available online: http://www.epa.gov/ttn/chief/trends/trends98/trends98.pdf, 2000b.

USEPA. *2002 Final National Emissions Inventory (NEI)*, available online: ftp://ftp.epa.gov/EmisInventory/2002finalnei/, 2006 (accessed October 10, 2006) [The 2002 NEI reports national annual emissions for total anthropogenic VOC emissions as 16.8 million tons, and total biogenic VOC emissions as 41.8 million tons].

4. WHAT WILL IT TAKE TO CLEAN THE AIR?

In this chapter we build on the conceptual description of fine particle formation and impacts in the MANE-VU region by looking at a typical fine particle pollution event and the meteorological and chemical conditions which contributed to its formation. As an illustration of how the conceptual elements laid out in Chapter 2 and 3 contribute to a pollution event under real-world circumstances, we examine a pollution event from 2002. We examine this event from two perspectives: (1) the broad spatial patterns of the formation and transport of particle air pollution and (2) the chronological sequence of events at a few discrete points where high temporal resolution monitoring was in place. We then proceed to examine likely emission reduction strategies that should be considered in light of the conceptual understanding of fine particle formation and transport developed in this report.

4.1. Meteorological and Pollution Overview of August 8-16, 2002

Annual and seasonal statistics are useful for understanding the general patterns of air pollution in our region, but it is also instructive to review specific high $PM_{2.5}$ episodes in order to shed more light on the meteorological circumstances under which high ambient concentrations of $PM_{2.5}$ are able to form from emitted precursor pollutants. Here we present an analysis of the high $PM_{2.5}$ and regional haze episode of August 2002 by reviewing surface maps from the period to provide a synoptic overview of major weather systems that were influencing air quality across the Northeast U.S. during that time.

Figure 4-1 through Figure 4-3, respectively, show eight-panel displays of afternoon fine particle concentrations as well as surface weather maps and back trajectories from 12Z (8 a.m. EDT) each day. The following chronology of events combines the meteorological insights with $PM_{2.5}$ concentration information to provide a basic storyline for analysis.

A slow-moving high pressure system centered over the Great Lakes set up northerly flow over MANE-VU on August 8. The high drifted southeast-ward and became extended over several days bringing high temperatures to the region. Calm conditions west of MANE-VU on August 10 were pivotal in the formation of fine aerosol concentrations, which began building in the Ohio River Valley. Over the next four days, concentrations in MANE-VU climbed into the 60-90 $\mu g/m^3$ range over a wide area before being swept out to sea by a series of frontal passages beginning on August 15.

8/8 – A high pressure system over the Great Lakes produces NW-N prevailing surface winds (~4-8 mph) throughout the region. Maximum daily temperatures approach or exceed 80º F.

8/9 – Wind speeds fall off but direction remains NW-N as the high moves into the central portion of MANE-VU. Temperatures rise as cloud cover declines.

8/10 – The high reaches the East Coast and stalls. Temperatures (except in northern-most areas) reach 90º F while surface-level winds turn to more southerly directions. Calm conditions through the morning hours in the lower Ohio River Valley promote creation of haze noted in surface observations.

8/11 – Circulation around the high (now near Cape Hatteras) becomes well established. Peak temperatures are in the low to mid-90's. Morning winds are light-to-calm in the area east of the Mississippi – the area of haze now reaches from Michigan to northern Texas and eastward to West Virginia and eastern Tennessee. A surface-level trough descends from north of the Great Lakes during the day, passes eastward through the Ohio River Valley and stalls over the Allegheny Mountains and southward.

8/12 – Temperatures exceed 90º F throughout MANE-VU except in coastal ME. The area of concentrated haze has pushed eastward and now extends from central ME to central PA. Haze builds throughout the day as circulation forces it to channel NE between the stalled trough and a cold front approaching from the Midwest.

8/13 – Calm conditions prevail as the trough reaches coastal NJ by 8 a.m. Generally clear skies allow temperatures to reach the mid-90's everywhere except in coastal ME. Dew points, which had been rising since 8/8, reach the upper 60's. Peak hourly fine aerosol concentrations are greater than 40 $\mu g/m^3$ everywhere in MANE-VU and exceed 90 $\mu g/m^3$ in some locations. By 8 p.m., showers associated with the approaching cold front have reached into Ohio.

8/14 – By 8 a.m. the trough has dissipated and the high is moving offshore. Dew points remain in the upper 60's and peak temperatures reach into the 90's everywhere and top 100 in several locations. Increased ventilation causes aerosol concentrations to drop throughout the day everywhere except ME where some locations peak above 60 $\mu g/m^3$ after midnight.

8/15 – The approaching cold front and associated showers fall apart during the morning hours. By 8 p.m., a new batch of moderate rain has intruded deeply into the region from the SW and has virtually pushed the haze out of the MANE-VU region.

8/16 – A new high building in over the upper Midwest pushes the remains of the showers out of the Northeast.

Figure 4-1. Spatially interpolated maps of fine particle concentrations August 9 – 16, 2002

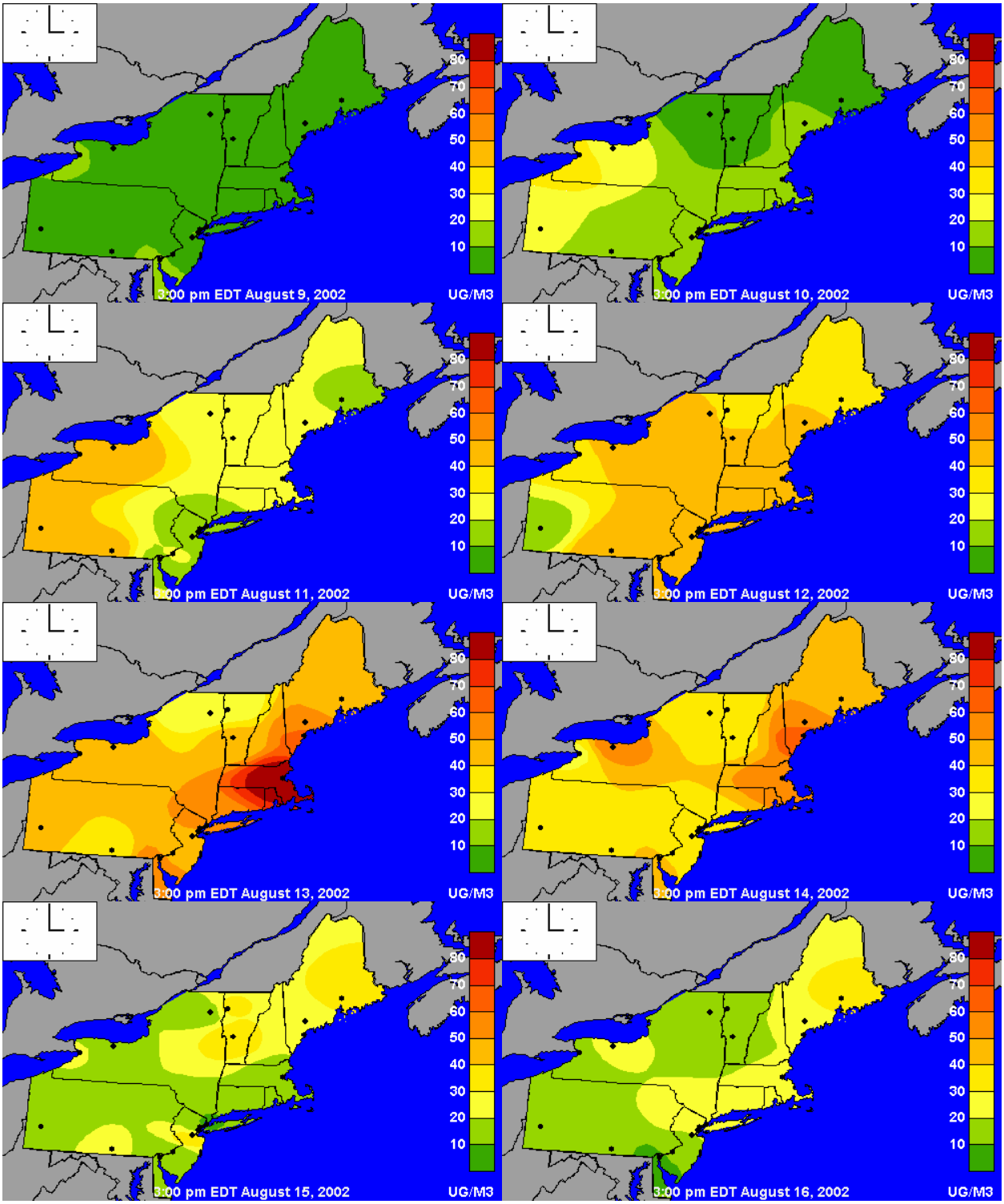


Figure 4-2. Surface weather maps for August 9-16, 2002

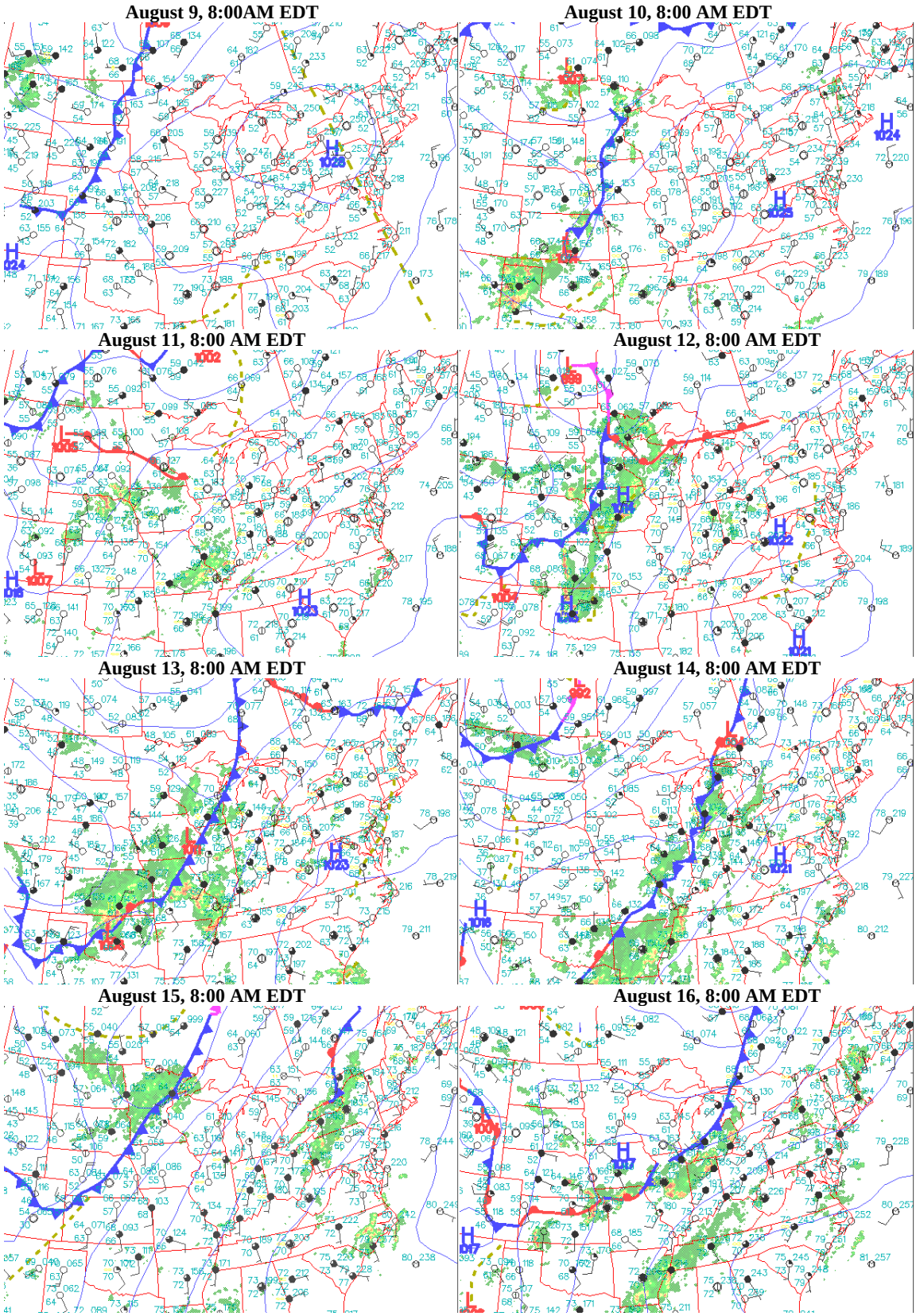
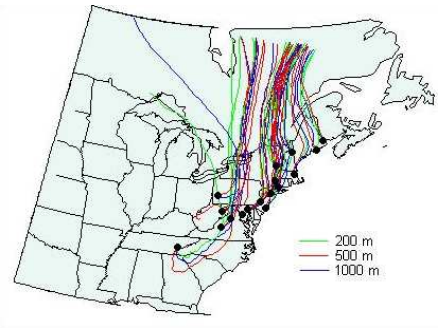
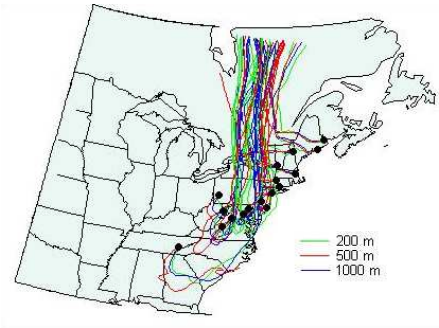


Figure 4-3. HYSPLIT 72-hour back trajectories for August 9-16, 2002

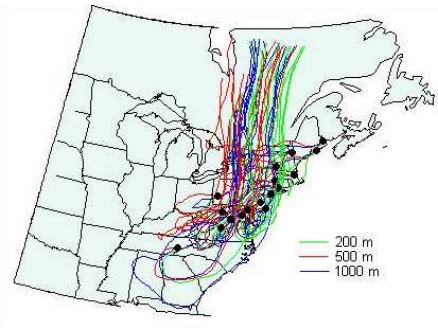
Aug 9, 2002 8 am EDT



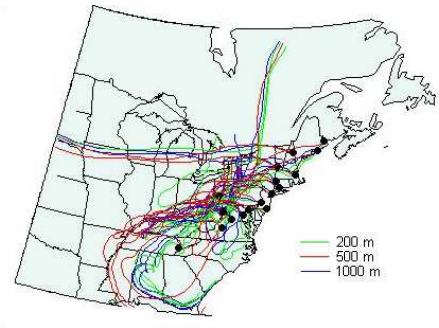
Aug 10, 2002 8 am EDT



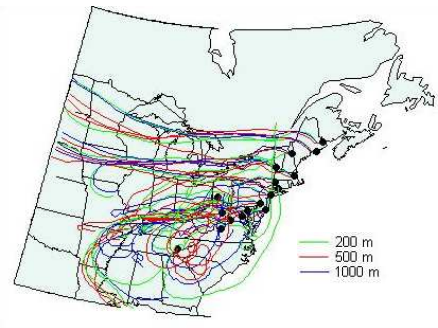
Aug 11, 2002 8 am EDT



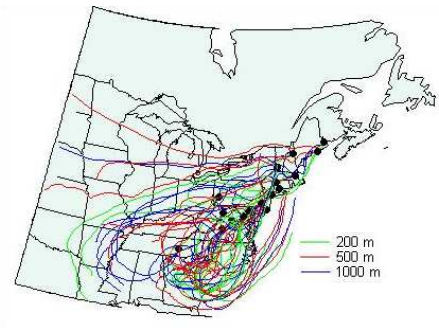
Aug 12, 2002 8 am EDT



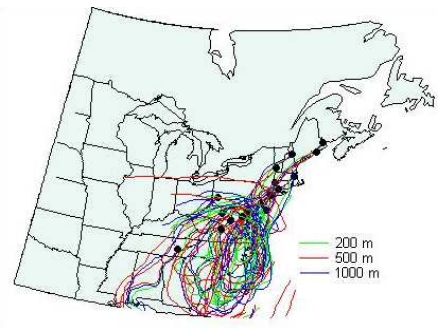
Aug 13, 2002 8 am EDT



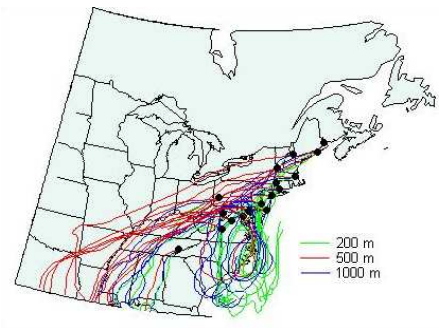
Aug 14, 2002 8 am EDT



Aug 15, 2002 8 am EDT



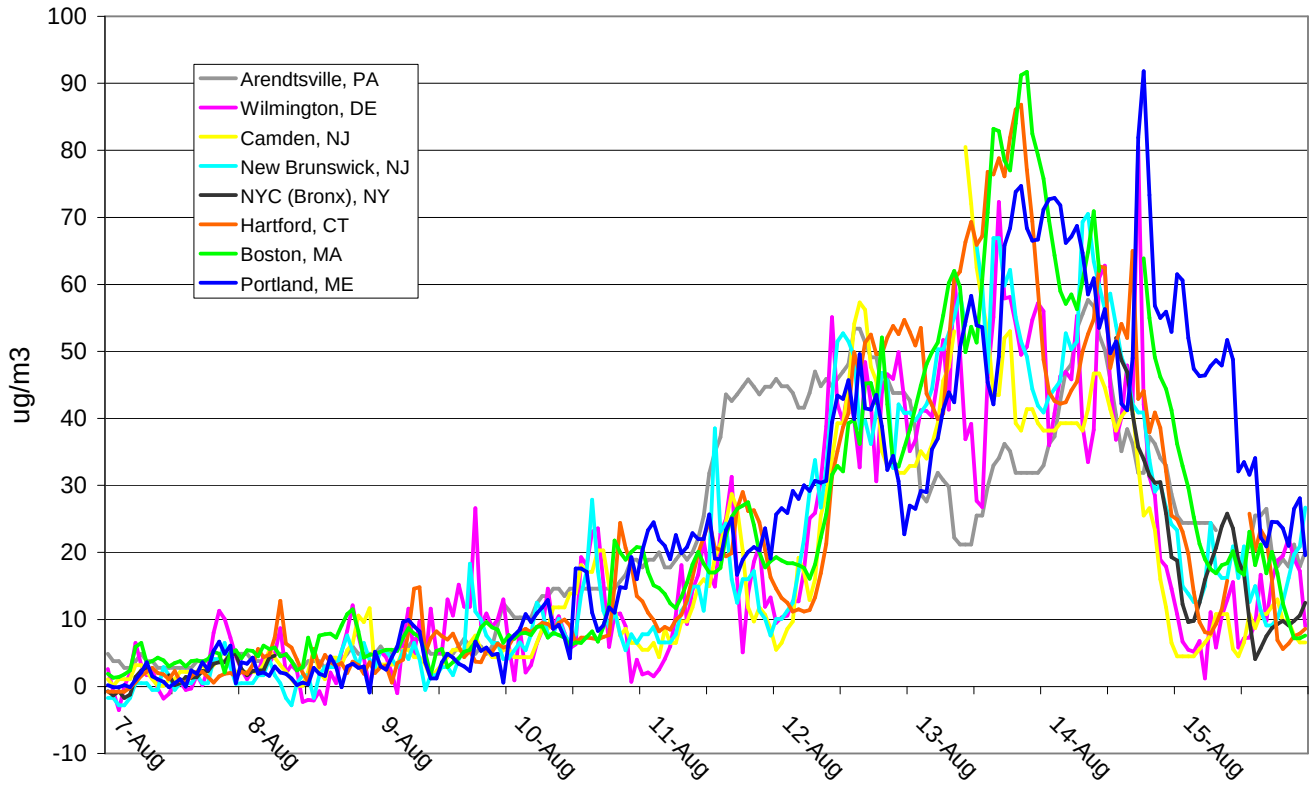
Aug 16, 2002 8 am EDT



4.2. Temporally and spatially resolved $PM_{2.5}$ measurements

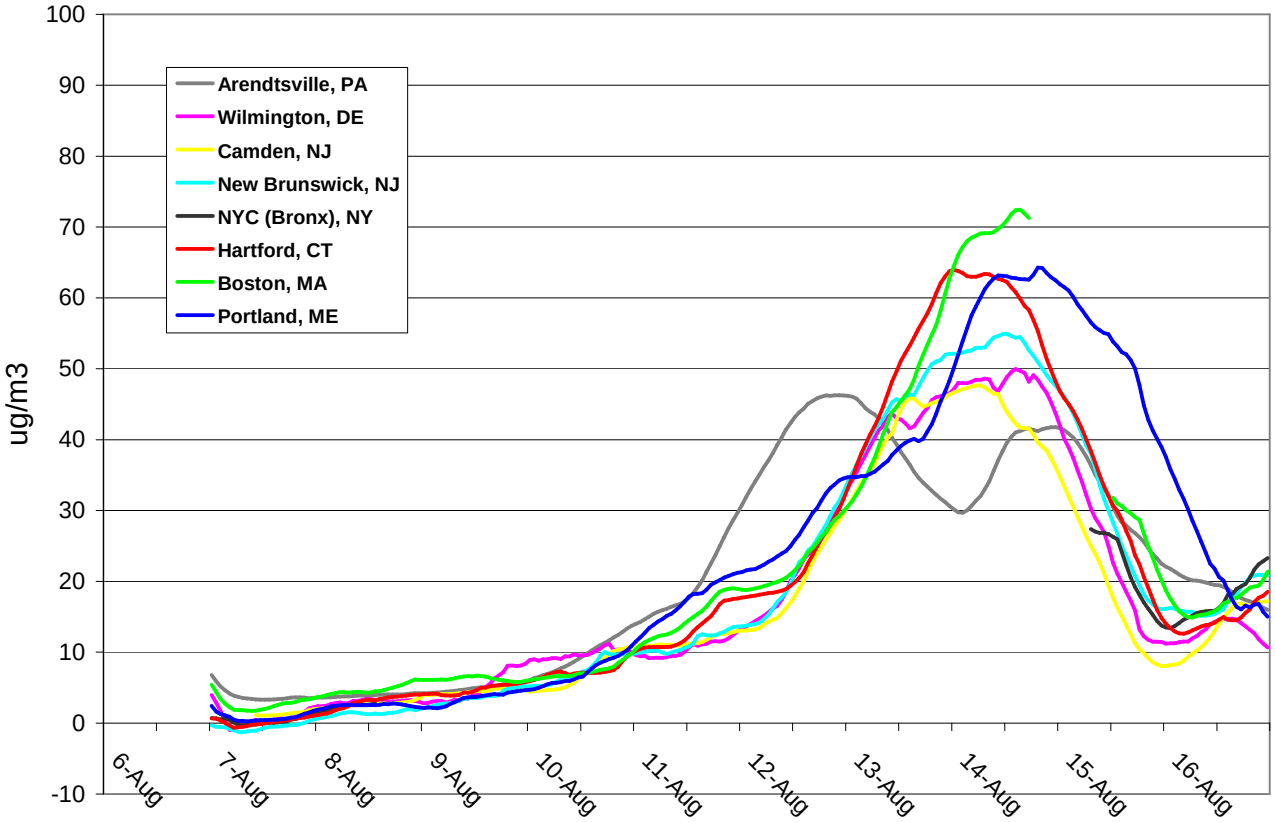
Higher temporal resolution data provide insight into how the events played out in much more detail than can be captured by eight frames on a page; however the most complete picture is obtained when these high *temporal* resolution data can be presented in the context of the relatively greater *spatial* detail provided by maps such as we have seen in Figure 4-1 through Figure 4-3. In Figure 4-4 and Figure 4-5, we present continuous $PM_{2.5}$ data (hourly average and 24-hour rolling average filtered, respectively) for the August 8-16, 2002 time period.

Figure 4-4. Hourly average fine aerosol at 8 sites during the August 2002 episode



Looking at Figure 4-4 in the context of the maps presented in the earlier figures, it is interesting to note the rapid increase, first, in Arendtsville, PA at noon on the 11th, followed by a rise in concentrations along the East Coast around noon on the 12th. This is consistent with Figure 4-1, which shows high $PM_{2.5}$ levels covering western Pennsylvania by 3 p.m. on the 11th and that high $PM_{2.5}$ area has moved over to cover the East Coast by 3 p.m. the next day. This also makes sense with respect to Figure 4-2 and Figure 4-3, which show the high pressure system established on the East Coast by the 11th with surface level back trajectories having shifted from northerly flow to slow southwesterly flow in the western portion of the domain by the morning of the 11th and the coastal sites having switched by the morning of the 12th.

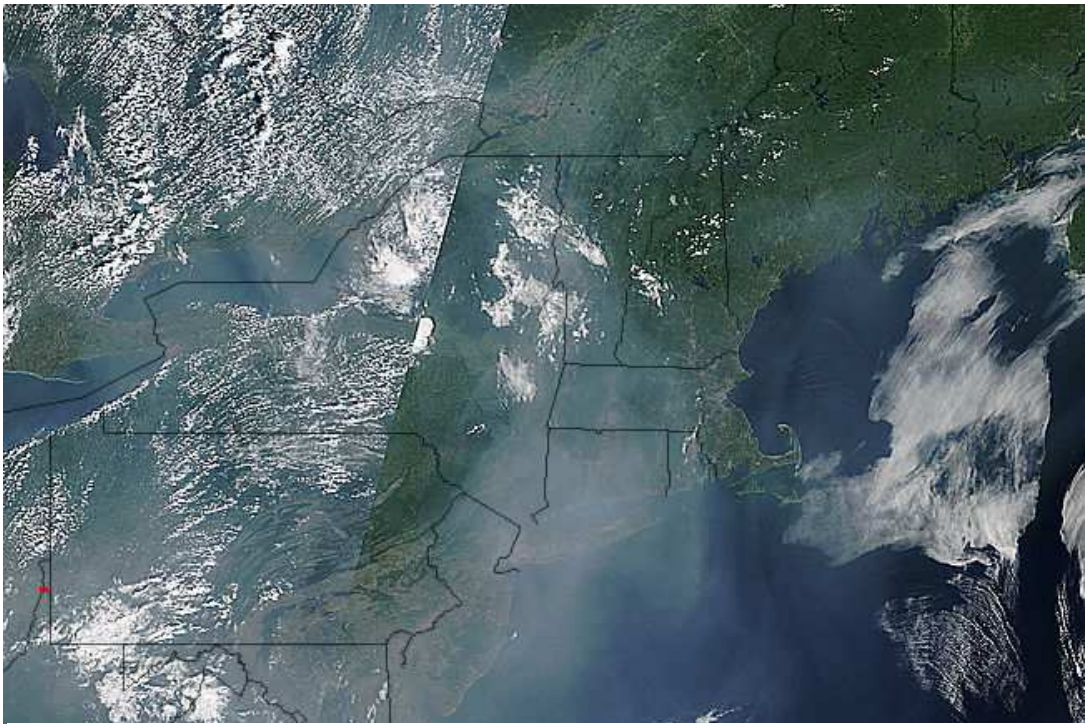
Figure 4-5. 24-hour rolling average fine aerosol at 8 MANE-VU sites during the August 2002 episode



Also note the very high levels observed close to mid-day on the 13th at sites between New York City and Portland, Maine. This is consistent with the strong gradients shown for 3 p.m. on the 13th in Figure 4-1. These rapid increases in concentration are easily explained by the back trajectories of Figure 4-3 that show the advancing front (at this point over Lake Michigan) beginning to push, at upper levels of the atmosphere, an air mass from the upper Midwest due east across the northern half of MANE-VU. At lower levels (see 200 meter trajectories), it can be seen that closer to the surface, this air mass had spent the previous three to four days winding around the Tennessee and Ohio River Valleys before it was driven into the northern reaches of MANE-VU at the peak of the pollution event.

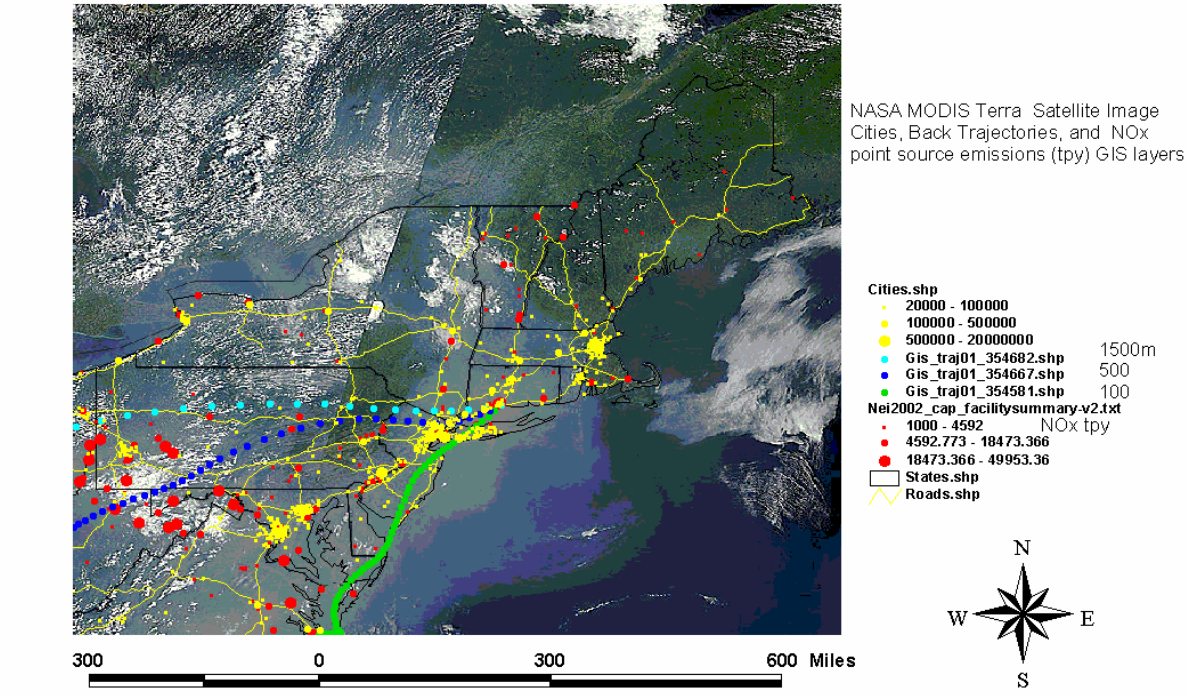
The following figures bring much of this information together in a single image. Figure 4-6 contains satellite photos from MODIS, a mosaic of two consecutive satellite passages on August 13, 2002 from NASA's TERRA satellite. Figure 4-7 shows the same image with geo-referenced activity data and inventory information layered on top to allow for simultaneous depiction of cities, roads, point source emissions, and back trajectories that play a role in the air pollution/haze that affected a large part of the Northeast during this episode.

Figure 4-6. Composite images from NASA's TERRA Satellite on August 13, 2002 showing fine particle pollution/haze.



Note the milky/gray haze due to particle pollution as distinct from the puffy white clouds over broad regions of southern New England and the eastern Mid-Atlantic region.

Figure 4-7. NASA MODIS Terra Satellite Image, Back Trajectories and NO_X Inventory



Geo-referenced activity and inventory data (on top of the satellite images presented above) demonstrating the relationship between observed pollution and upper level winds (driving weather patterns from West to East), mid-level winds (tracking back to major point sources), and lower level winds (tracking back to major population centers along the East Coast).

4.3. Implications for control strategies

A 2003 assessment of fine particulate matter by NARSTO[15] states, "[c]urrent air-quality management approaches focusing on reductions of emissions of SO_2, NO_X, and VOCs are anticipated to be effective first steps towards reducing $PM_{2.5}$ across North America, noting that in parts of California and some eastern urban areas VOC (volatile organic compounds) emissions could be important to nitrate formation."

This conclusion seems to be well supported by the historical record which documents a pronounced decline in particulate sulfate concentrations across the eastern United States during the 1990s. The timing of this observed decline suggests that this is linked to reductions in SO_2 emissions resulting from controls implemented under the federal Acid Rain program beginning in the early to mid-1990s. From 1989 to 1998, SO_2 emissions in the eastern half of the country — that is, including all states within a region defined by the western borders of Minnesota and Louisiana — declined by about 25 percent. This decline in SO_2 emissions correlated with a decline of about 40 percent in average SO_2 and sulfate concentrations, as measured at Clean Air States and Trend Networks (CASTNet) monitoring sites in the same region over the same time period. In fact, at prevailing levels of atmospheric SO_2 loading, the magnitudes of the emissions and concentration changes were not statistically different. This finding suggests that regional reductions in SO_2 emissions have produced near-proportional reductions of particulate sulfate in the eastern United States (NARSTO, 2003). Reductions since 1990 in precursor SO_2 emissions are likely also responsible for a continued decline in median sulfate concentrations in the northeastern United States. Nevertheless, episodes of high ambient sulfate concentrations (with peak levels well above the regional median or average) continue to occur, especially during the summertime when regional transport from the Ohio River Valley is also at its peak. This suggests that further reductions in regional and local SO_2 emissions would provide significant further air quality and visibility benefits (NARSTO, 2003).

For urban areas of the eastern United States, an effective emissions management approach may be to combine regional SO_2 control efforts aimed at reducing summertime $PM_{2.5}$ concentrations with local SO_2 and OC control efforts. Local SO_2 reductions would help reduce wintertime $PM_{2.5}$ concentrations, while OC reductions can help reduce overall $PM_{2.5}$ concentrations year-round. For areas with high wintertime $PM_{2.5}$ levels, strategies that involve NO_X reductions may also be effective (NARSTO, 2003).

Further support for this general approach may be found in a review of several studies by Watson (2002) which concluded that SO_2 emission reductions have in most cases been accompanied by statistically significant reductions in ambient sulfate concentrations. One study (Husar and Wilson, 1993) shows that regionally averaged light extinction closely tracks regionally averaged SO_2 emissions for the eastern United States from 1940 through the mid-1980s. Another study by Malm *et al.* (2002) shows that

[15] NARSTO was formerly an acronym for the "North American Research Strategy for Tropospheric Ozone." More recently, the term NARSTO became simply a wordmark signifying a tri-national, public-private partnership for dealing with multiple features of tropospheric pollution, including ozone and suspended particulate matter. For more information on NARSTO see http://www.cgenv.com/Narsto/.

regionally averaged emissions and ambient concentrations decreased together from 1988 through 1999 over a broad region encompassing the states of Connecticut, Delaware, Illinois, Indiana, Kentucky, Maine, Massachusetts, Maryland, Michigan, New Hampshire, New Jersey, New York, Ohio, Pennsylvania, Rhode Island, Vermont, Virginia, Wisconsin, and West Virginia (Watson, 2002).

These studies and available data from the IMPROVE (Interagency Monitoring of Protected Visual Environment) monitoring network provide strong evidence that regional SO_2 reductions have yielded, and will continue to yield, reductions in ambient secondary sulfate levels with subsequent reductions in regional haze and associated light extinction. They indicate that reductions in anthropogenic primary particle emissions will also result in visibility improvements, but that these will not have a zone of influence as large as those of the secondary aerosols (Watson, 2002).

Watson (2002) notes that during the 65 years in which the regional haze program aims to reach its final visibility goals, several opportunities to revise this basic control approach will arise through the decadal SIP cycle. This enables new scientific results to continue to exert a positive influence as states implement new regulatory control programs for SO_2, NO_X and VOCs, and as ambient concentrations of these pollutants change relative to each other and relative to ambient ammonia levels. As these relationships between species change, atmospheric chemistry may dictate a revised control approach to those previously described. Further research on these issues should be a priority for supporting 2018 SIP submissions. They include the possibility that:

- Reduction of sulfate in a fully neutralized atmosphere (excess ammonia) could encourage ammonium nitrate formation.
- Ever-greater emissions reductions could be required to produce a given level of improvement in ambient pollutant concentrations because of non-linearities in the atmospheric formation of sulfate.
- Changes in ambient conditions favoring the aqueous oxidation of sulfate (this pathway largely accounts for the non-linearity noted above) may have implications for future emissions control programs. Causes of changing ambient conditions could include, for example, climate change.

West *et al.* (1999) examine a scenario for the eastern United States where $PM_{2.5}$ mass decreases linearly with ammonium sulfate until the latter is fully neutralized by ammonia. Further reductions would free ammonia for combination with gaseous nitric acid that, in turn, would slightly increase $PM_{2.5}$ until all of the nitric acid is neutralized and further sulfate reductions are reflected in lower $PM_{2.5}$ mass. This is an extreme case that is more relevant to source areas (e.g., Ohio) where nitric acid (HNO_3) is more abundant than in areas with lower emissions (e.g., Vermont) (Watson, 2002).

In most situations with non-neutralized sulfate (typical of the eastern United States), ammonia is a limiting agent for the formation of nitrate but will not make any difference until sulfate is reduced to the point where it is completely neutralized. At that point, identifying large sources of ammonia emissions will be important. This point is likely to be many years in the future, however (Watson, 2002).

Based on analyses using the Community Multi-Scale Air Quality (CMAQ) model, the aqueous phase production of sulfate in the Northeast appears to be very oxidant limited and hence non-linear. Thus, conditions that are conducive to a dominance of the gas-phase production pathway drive the summer peaks in ambient sulfate levels. Nonetheless, the expected reduction in ambient sulfate levels resulting from a given reduction in SO_2 emissions is less than proportional overall due to the non-linearity introduced by the aqueous pathway for sulfate formation (NARSTO, 2003). These non-linearity effects are more pronounced for haze than for sulfate deposition, especially at higher sulfate air concentrations (USNPS, 2003).

Finally, we note that because visibility in the clearest areas is sensitive to even minute increases in particle concentrations, strategies to preserve visibility on the clearest days may require stringent limits on emissions growth. In this context, even the dilute emissions from distant sources can be important (NARSTO, 2003)

4.4. Conclusion: Simplifying a complex problem

A conceptual understanding of fine particles from a regional perspective across MANE-VU and throughout the eastern U.S. is well understood, yet remains complex due to the multiplicity of source regions (both regional and local), pollutants (SO_2, NO_X, organic carbon, and primary $PM_{2.5}$), and seasons (summer and winter) that are involved in fine particle formation.

Regional approaches to the control of precursor SO_2 and NO_X emissions have been started through Title IV of the Clean Air Act, the NO_X SIP Call, the CAIR, and the establishment and support of Regional Planning Organizations to assist with Regional Haze Rule compliance. With the modeling foundation developed for the CAIR program, the USEPA has presented a compelling technical case on the need for additional regional SO_2 and NO_X reductions in the eastern U.S. to reduce particulate levels and protect public health. While states in the Northeast disagree with the extent of SO_2 and NO_X reductions and the timeline for those reductions to occur, the program is an excellent next step toward reducing fine particles in MANE-VU. It is tempting to suggest that the regional control of SO_2 and NO_X are the extent of the problem facing MANE-VU, but as the conceptual description contained in this report demonstrates, the reduction of fine particles in the eastern U.S. requires a careful balance of regional and local controls for SO_2, NO_X, sources of organic carbon and primary $PM_{2.5}$ during both summer and winter.

The (relatively) higher emissions of SO_2 and NO_X from regions upwind of MANE-VU as well as the long "reach" of sulfate pollution requires continued regional control of these fine particle precursors. However, local accumulation of SO_2-derived sulfate, NO_X-derived nitrate, and primary PM (mostly in the form of black carbon/diesel exhaust) can significantly boost urban $PM_{2.5}$ levels. Residential wood combustion in rural river valleys can significantly raise PM levels as well and affect rural visibility in areas near to Class I areas.

The balance between regional and local controls parallels the balance that needs to be achieved between pollutants. The regional contribution to fine particle pollution is driven by sulfates and organic carbon, whereas the local contribution to $PM_{2.5}$ is derived

from SO_2, NO_X, organic carbon, and primary $PM_{2.5}$ (including black carbon/diesel exhaust).

Finally, control strategies which focus on regional SO_2 emissions reductions are needed throughout the summer and winter months, suggesting that a year-round approach to control is needed. Urban nonattainment counties with local emissions of NO_X and VOC will be driven to reduce these emissions during the summer for ozone benefits, but these same pollutants – as well as primary particulate emissions – contribute to high $PM_{2.5}$ levels in winter, suggesting that annual controls for all of these pollutants make sense in a multi-pollutant context. Finally, residential wood smoke near Class I areas is clearly a winter-only issue, and further controls may be desirable near specific Class I sites where organic carbon is a contributor on the 20 percent worst visibility days that occur in winter months.

To bring attainment to the current fine particle nonattainment counties and meet reasonable progress goals toward national visibility goals, there continues to be a need for more regional SO_2 and NO_X reductions coupled with appropriate local SO_2, NO_X, VOC, and primary $PM_{2.5}$ (including diesel exhaust) controls where local accumulation is shown to add to the regional burden of sulfate and nitrate $PM_{2.5}$ (primarily in winter). These local controls will vary by location and by season, but the regional control of SO_2 and NO_X should be maintained on an annual basis given the contribution of regional sulfate and nitrate to fine particle peaks during both summer and winter months.

Appendix A: Excerpts from EPA Guidance Document, Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, $PM_{2.5}$, and Regional Haze

APPENDIX A: EPA GUIDANCE DOCUMENT EXERPT

11.0 How Do I Get Started? - A "Conceptual Description"

A State/Tribe should start developing information to support a modeled attainment demonstration by assembling and reviewing available air quality, emissions and meteorological data. Baseline design values should be calculated at each monitoring site, as described in Section 3. For PM applications, speciated data should be reviewed to get a sense of what component(s) might be contributing most significantly to nonattainment or light extinction. If past modeling has been performed, the emission scenarios examined and air quality predictions may also be useful. Readily available information should be used by a State/Tribe to develop an initial conceptual description of the nonattainment or reasonable haze problem in the area which is the focus of a modeled demonstration. A conceptual description is instrumental for identifying potential stakeholders and for developing a modeling/analysis protocol. It may also influence a State's choice of air quality model, modeling domain, grid cell size, priorities for quality assuring and refining emissions estimates, and the choice of initial diagnostic tests to identify potentially effective control strategies. In general, a conceptual description is useful for helping a State/Tribe identify priorities and allocate resources in performing a modeled demonstration.

In this Section, we identify key parts of a conceptual description. We then present examples of analyses which could be used to describe each of these parts. We note that initial analyses may be complemented later by additional efforts performed by those implementing the protocol.

11.1 What Is A "Conceptual Description"?

A "conceptual description" is a qualitative way of characterizing the nature of an area's nonattainment or regional haze problem. It is best described by identifying key components of a description. Examples are listed below. There are 3 different examples. One each for ozone, annual $PM_{2.5}$, and regional haze. The examples are not necessarily comprehensive. There could be other features of an area's problem which are important in particular cases. For purposes of illustration later in the discussion, we have answered each of the questions posed below. Our responses appear in parentheses.

11.1.1 8-Hour Ozone NAAQS

1. Is the nonattainment problem primarily a local one, or are regional factors important?
(Surface measurements suggest transport of ozone close to 84 ppb is likely. There are some other nonattainment areas not too far distant.)

2. Are ozone and/or precursor concentrations aloft also high?
(There are no such measurements.)

3. Do violations of the NAAQS occur at several monitoring sites throughout the nonattainment area, or are they confined to one or a small number of sites in proximity to one another?
(Violations occur at a limited number of sites, located throughout the area.)

4. Do observed 8-hour daily maximum ozone concentrations exceed 84 ppb frequently or just on a few occasions?
(This varies among the monitors from 4 times up to 12 times per year.)

5. When 8-hour daily maxima in excess of 84 ppb occur, is there an accompanying characteristic spatial pattern, or is there a variety of spatial patterns?
(A variety of patterns is seen.)

6. Do monitored violations occur at locations subject to mesoscale wind patterns (e.g., at a coastline) which may differ from the general wind flow?
(No.)

7. Have there been any recent major changes in emissions of VOC or NO_X in or near the nonattainment area? If so, what changes have occurred?
(Yes, several local measures [include a list] believed to result in major reductions in VOC [quantify in tons per summer day] have been implemented in the last five years. Additionally, the area has seen large regional NO_X reductions from the NO_X SIP call.)

8. Are there discernible trends in design values or other air quality indicators which have accompanied a change in emissions?
(Yes, design values have decreased by about 10% at four sites over the past [x] years. Smaller or no reductions are seen at three other sites.)

9. Is there any apparent spatial pattern to the trends in design values?
(No.)

10. Have ambient precursor concentrations or measured VOC species profiles changed?
(There are no measurements.)

11. What past modeling has been performed and what do the results suggest?
(A regional modeling analysis has been performed. Two emission scenarios were modeled: current emissions and a substantial reduction in NO_X emissions throughout the regional domain. Reduced NO_X emissions led to substantial predicted reductions in 8-hour daily maximum ozone in most locations, but changes near the most populated area in the nonattainment area in question were small or nonexistent.)

12. Are there any distinctive meteorological measurements at the surface or aloft which appear to coincide with occasions with 8-hour daily maxima greater than 84 ppb?
(Other than routine soundings taken twice per day, there are no measurements aloft. There is no obvious correspondence with meteorological measurements other than daily maximum temperatures are always > 85 F on these days.)

Using responses to the preceding questions in this example, it is possible to construct an initial conceptual description of the nonattainment area's ozone problem. First, responses to questions 1 and 11 suggest there is a significant regional component to the area's nonattainment problem. Second, responses to questions 3, 4, 7, 8, and 11 indicate there is an important local component to the area's nonattainment problem. The responses to questions 4, 5 and 12 indicate that high ozone concentrations may be observed under several sets of meteorological conditions. The responses to questions 7, 8, and 11 suggest that ozone in and near the nonattainment area may be responsive to both VOC and NO_X controls and that the extent of this response may vary spatially. The response to question 6 suggests that it may be appropriate to develop a strategy using a model with 12 km grid cells.

The preceding conceptual description implies that the State/Tribe containing the nonattainment area in this example will need to involve stakeholders from other, nearby States/Tribes to develop and implement a modeling/analysis protocol. It also suggests that a nested regional modeling analysis will be needed to address the problem. Further, it may be necessary to model at least several distinctive types of episodes and additional analyses will be needed to select episodes. Finally, sensitivity (i.e., diagnostic) tests, or other modeling probing tools, will be needed to assess the effects of reducing VOC and NO_X emissions separately and at the same time.

11.1.2 Annual $PM_{2.5}$ NAAQS

1. Is the nonattainment problem primarily a local one, or are regional factors important?
(Surface measurements suggest that only design values in or immediately downwind of the city violate the NAAQS. However, other nearby design values come close to the concentration specified in the NAAQS)

2. What is the relative importance of measured primary and secondary components of $PM_{2.5}$ measured at sites violating the NAAQS?
(Secondary components (i.e., SO_4, NO_3, OC) constitute about 80% of the measured mass of $PM_{2.5}$. There are higher concentrations of primary $PM_{2.5}$ in the core urban area compared to the suburbs and more rural areas.)

3. What are the most prevalent components of measured $PM_{2.5}$?
(The most important components in ranked order are mass associated with SO_4, OC and inorganic primary particulate matter (IP)).

4. Does the measured mix of PM components appear to roughly agree with mix of emission categories surrounding the monitoring sites?
(No. Relative importance of measured crustal material (IP) appears less than what might be inferred from the inventory).

5. Do there appear to be any areas with large gradients of primary $PM_{2.5}$ in monitored or unmonitored areas?
(Cannot really tell for sources of crustal material until we resolve the preceding

inventory/monitoring discrepancy. There are no other obvious major sources of primary particulate matter).

6. Is there any indication of what precursor might be limiting formation of secondary particulate matter?
(No indicator species analyses have been performed. Past analyses performed for ozone-related SIP revisions suggest that ozone in this area may be limited by availability of VOC).

7. Do monitored violations occur at locations subject to mesoscale wind patterns (e.g., at a coastline) which may differ from the general wind flow?
(No.)

8. Have there been any recent major changes in emissions of PM or its precursors in or near the nonattainment area? What?
(Yes, measures believed to result in major reductions in VOC and NO_X have been implemented in the last 5 years. Reductions in power plant NO_X have resulted from the NO_X SIP call and SO_2 emissions reductions have resulted from the national program to reduce acid deposition.)

9. Are there discernible trends in design values or other air quality indicators which have accompanied a change in emissions?
(The trend appears to be downward, but the most recent air quality data has been higher. Overall, the period of record is insufficiently long to tell).

10. Is there any apparent spatial pattern to the trends in design values?
(No.)

11. What past modeling has been performed and what do the results suggest?
(A regional modeling analysis has been performed for ozone and $PM_{2.5}$. Two emission scenarios were modeled: current emissions and a substantial reduction in NO_X and SO_2 emissions throughout a regional domain. Reduced NO_X emissions led to substantial predicted reductions in 8-hour daily maximum ozone in most locations. Modeled SO_2 reductions from the CAIR rule had a strong impact on sulfate concentrations)

12. Are there any distinctive meteorological measurements at the surface or aloft which appear to coincide with occasions with $PM_{2.5}$ concentrations in excess of 15.0 :g/m3?
(Other than routine soundings taken twice per day, there are no measurements aloft. There is no obvious correspondence with meteorological measurements other than daily maximum temperatures are often > 85F on days with the highest $PM_{2.5}$ observations.)

13. Do periods with high measured particulate matter or components of particulate matter appear to track each other or any other measured pollutant?
(There appears to be some correspondence between measured high concentrations of SO_4 and ozone).

Using responses to the preceding questions in this example, it is possible to construct an initial conceptual description of the nonattainment area's ozone problem. First, responses to questions 1, 2 and 3 suggest there is a significant regional component to the area's nonattainment problem. Second, responses to questions 1 and 3 indicate there is a local component to the problem. The responses to questions 11,12 and 13 suggest that there may be a link between reducing ozone and reducing particulate matter. Thus, it may be appropriate to assess effects of previously committed to strategies to reduce ozone and national PM control measures before simulating additional control measures. The responses to questions 4 and 5 suggest that it is premature to determine whether a "local area analysis" will be needed. The response to question 7 suggests that it may not be necessary to model with very small grid cells, at least for the secondary components of $PM_{2.5}$.

The preceding conceptual description implies that the State containing the nonattainment area in this example will need to involve stakeholders from other, nearby States to develop and implement a modeling/analysis protocol. It also suggests that a nested regional modeling analysis will be needed to address the problem.

11.1.3 Example reasonable progress application

1. What components of particulate matter appear to have high concentrations on days with poor visibility?
(Mass associated with SO_4 and coarse particulate matter (CM) seem to have the highest concentrations on most such days).

2. What are typical values for the humidity adjustment factor during the times of year when most of the days with poor visibility occur?
(Typical values appear to be about "4.0").

3. Does visibility appear to track well among nearby Class I areas?
(Yes, but not always).

4. Does poor visibility seem to occur under any specific meteorological conditions?
(This information is not readily available).

5. Does poor visibility seem to coincide with high observed concentrations of any particular other pollutant?
(There seems to be some correspondence with high regional ozone concentrations)

6. What components of particulate matter appear to have relatively high concentrations on days with good visibility?
(Coarse particulate matter and OC)

7. What are typical values for the humidity adjustment factor during times of year when most of the days with good visibility occur?
(About "2.3")

8. Does good visibility appear to occur under any specific meteorological conditions? (Don't know.)

Answers to the preceding questions suggest that strategies to reduce sulfate concentrations and, perhaps, regional ozone concentrations might be effective in reducing light extinction on days when visibility is currently poor. The responses suggest that a strategy which focuses on this alone should first be tried for the days with good visibility as well. Even though sulfate concentrations appear low on such days, the fact that sulfates scatter light efficiently (see Equation (6.1)) and relative humidity is still high enough to enhance this effect is worth considering. Responses suggest that further meteorological analyses would be worthwhile prior to selecting strategies to simulate with a resource intensive regional model.

It should be clear from the preceding examples that the initial conceptual description of an area's nonattainment problem draws on readily available information and need not be detailed. It is intended to help launch development and implementation of a modeling/analysis protocol in a productive direction. It will likely be supplemented by subsequent, more extensive modeling and ambient analyses performed by or for those implementing the modeling/analysis protocol discussed in Section 12.0.

Questions like those posed in Section 11.1 can be addressed using a variety of analyses ranging in complexity from an inspection of air quality data to sophisticated mathematical analyses. We anticipate the simpler analyses will often be used to develop the initial conceptual description. These will be followed by more complex approaches or by approaches requiring more extensive data bases as the need later becomes apparent. These analyses are intended to channel resources available to support modeled attainment demonstrations onto the most productive paths possible. They will also provide other pieces of information which can be used to reinforce conclusions reached with an air quality model, or cause a reassessment of assumptions made previously in applying the model. As noted in Section 7, corroboratory analyses should be used to help assess whether a simulated control strategy is sufficient to meet the NAAQS.

Appendix B: Monitoring Data from Class I sites in MANE-VU

Below are figures that were developed by Tom Downs of the Maine Department of Environmental Protection. These figures represent baseline monitoring data for the Class I sites (and Washington DC) based on IMPROVE monitoring network data using the EPA approved "default" algorithm for calculating reconstructed extinction and estimating natural background conditions. These statistics may need to be recreated using the alternative methodology approved by the IMPROVE steering committee and adopted by the MANE-VU states. Glide path graphs were created on the VIEWS website (http://vista.cira.colostate.edu/views/) using the Annual Summary Trends tool. Seasonal graphs were created from data downloaded from the VIEWS website using the Annual Summary Composition tool and should be updated to include 2004 data for a complete description of regional haze baseline data.

APPENDIX B: MONITORING DATA FROM CLASS I SITES IN MANE-VU

Figure B-1. Monitoring Data from Acadia NP, ME

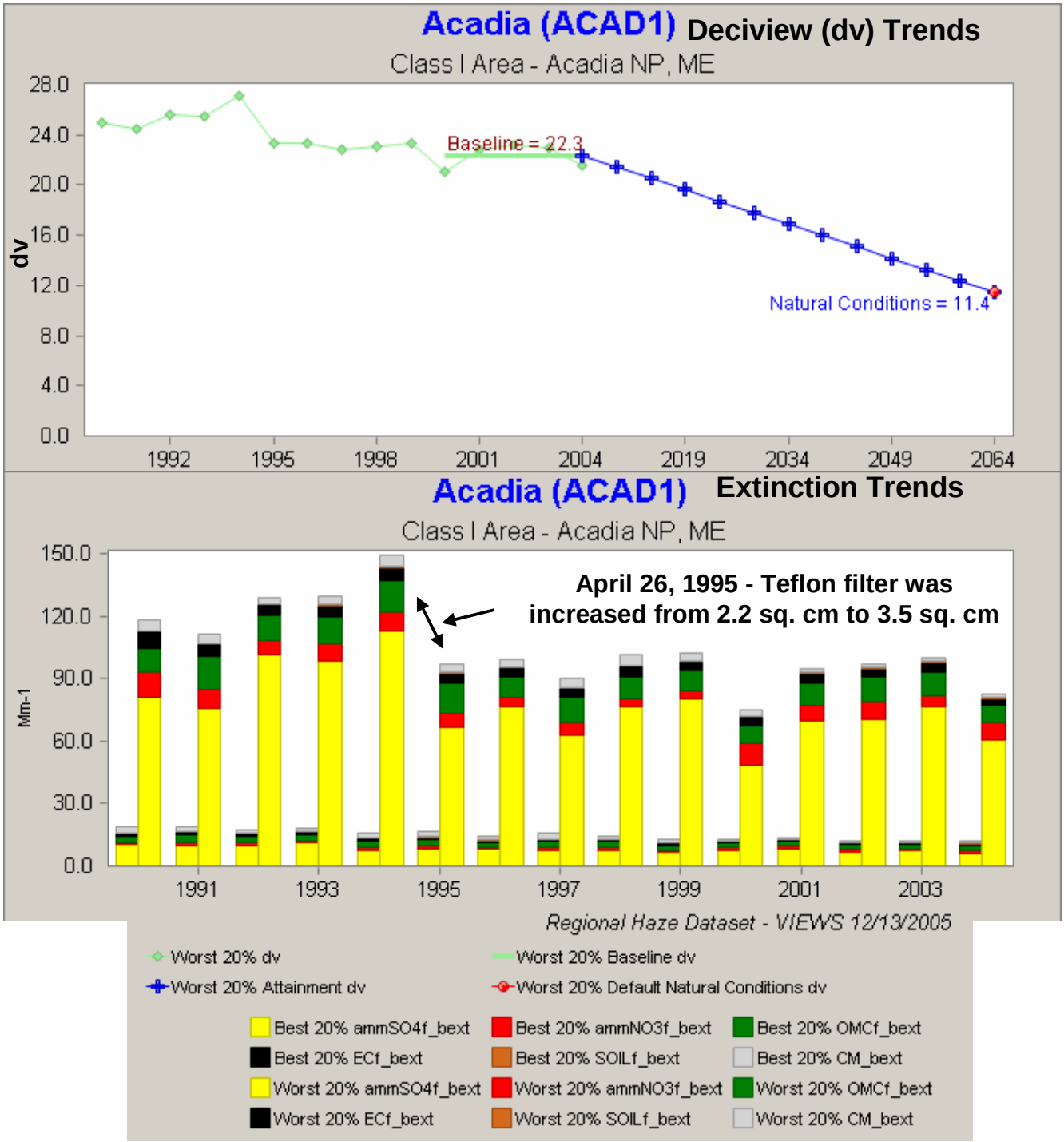


Figure B-2. Monitoring Data from Brigantine, ME

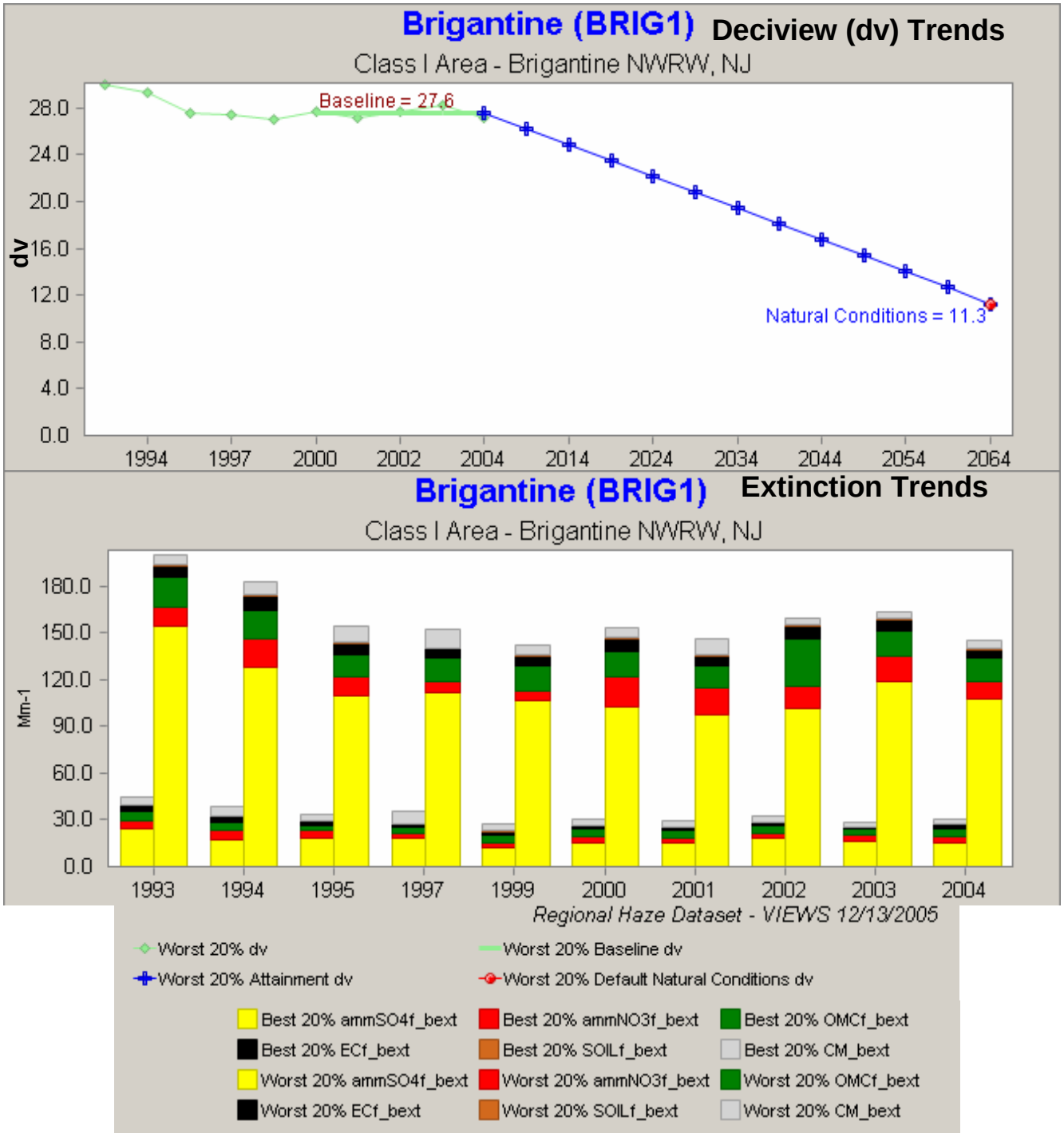


Figure B-3. Monitoring Data from Great Gulf, NH

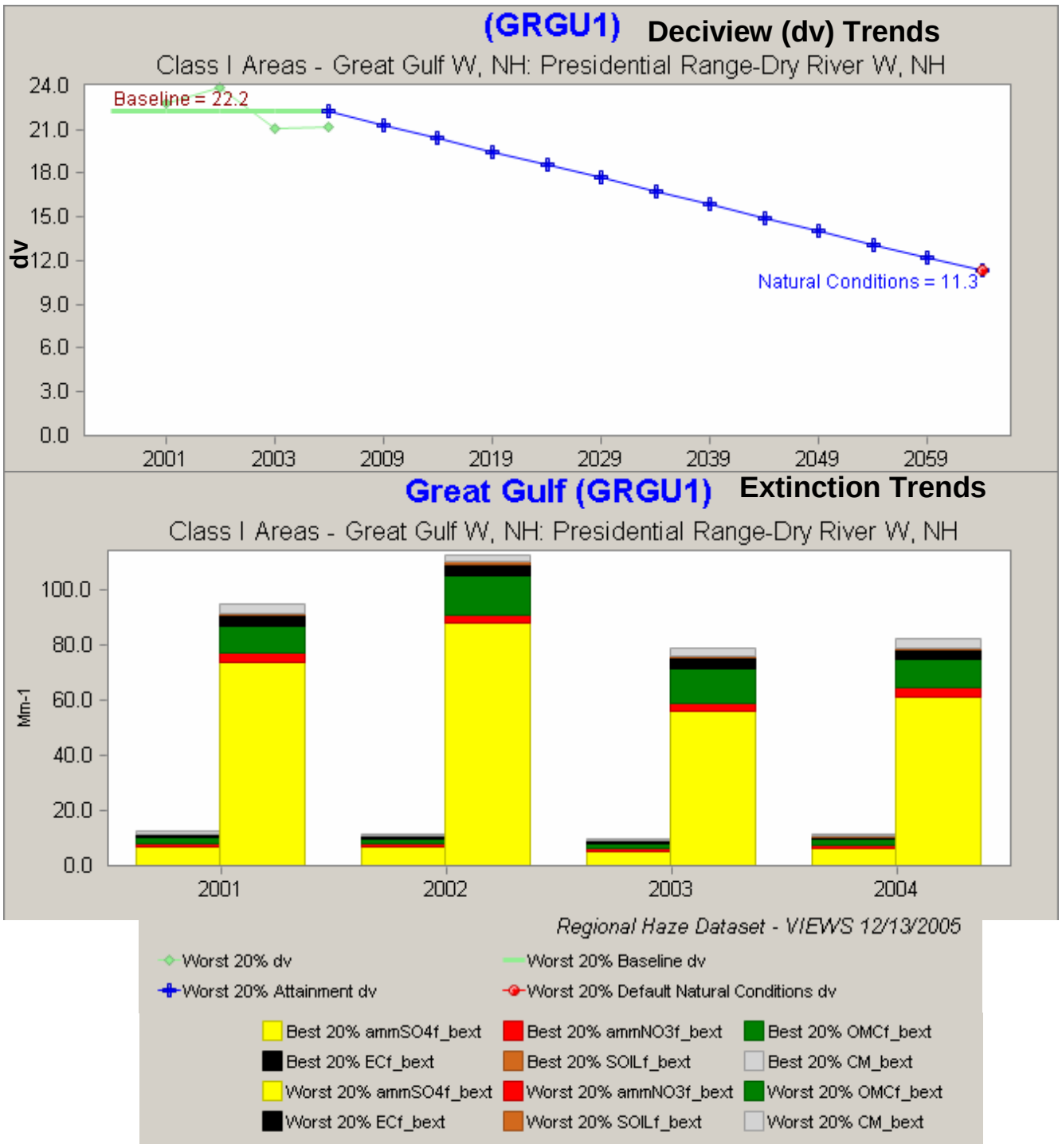


Figure B-4. Monitoring Data from Lye Brook, VT

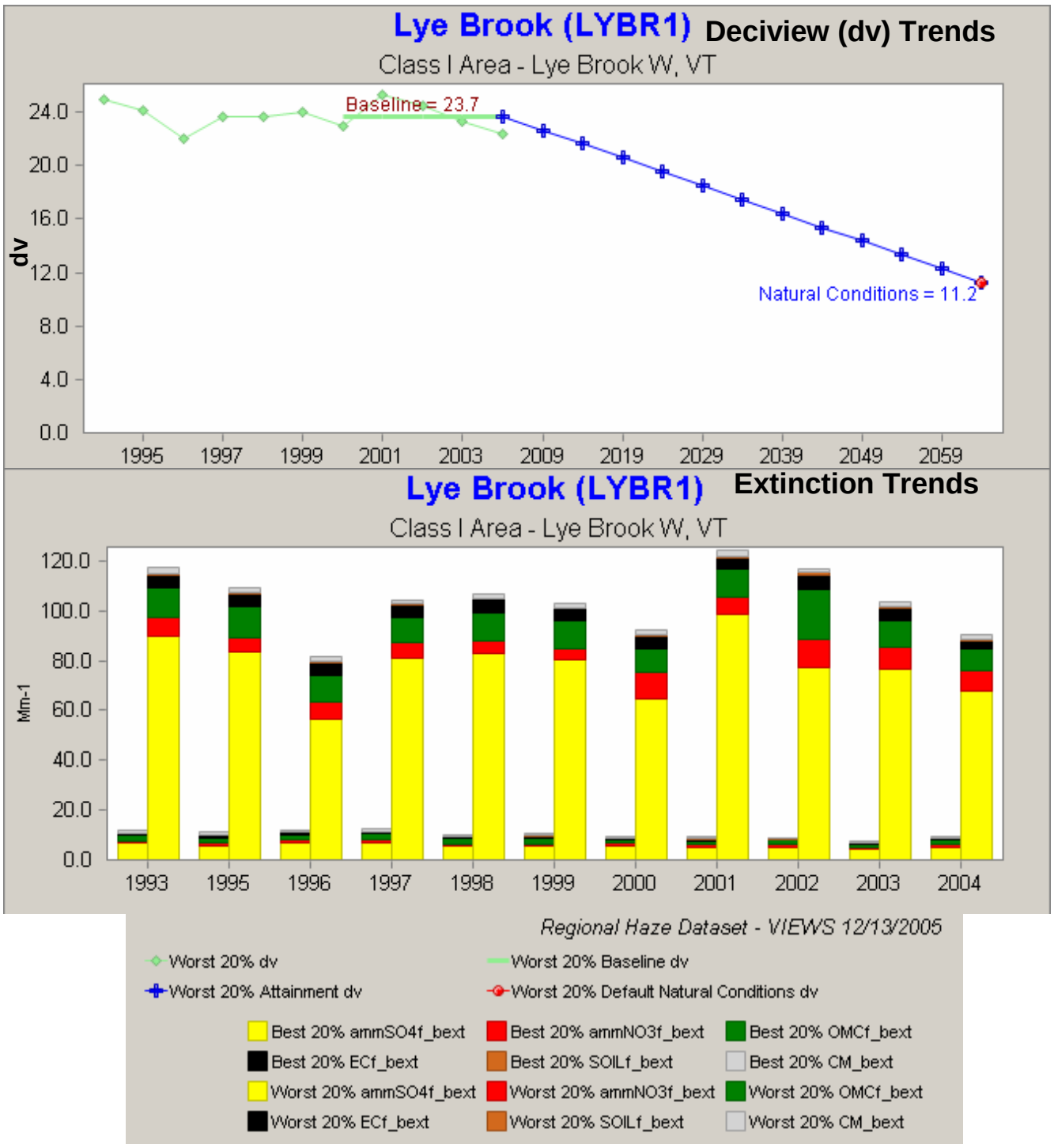


Figure B-5. Monitoring Data from Moosehorn, ME

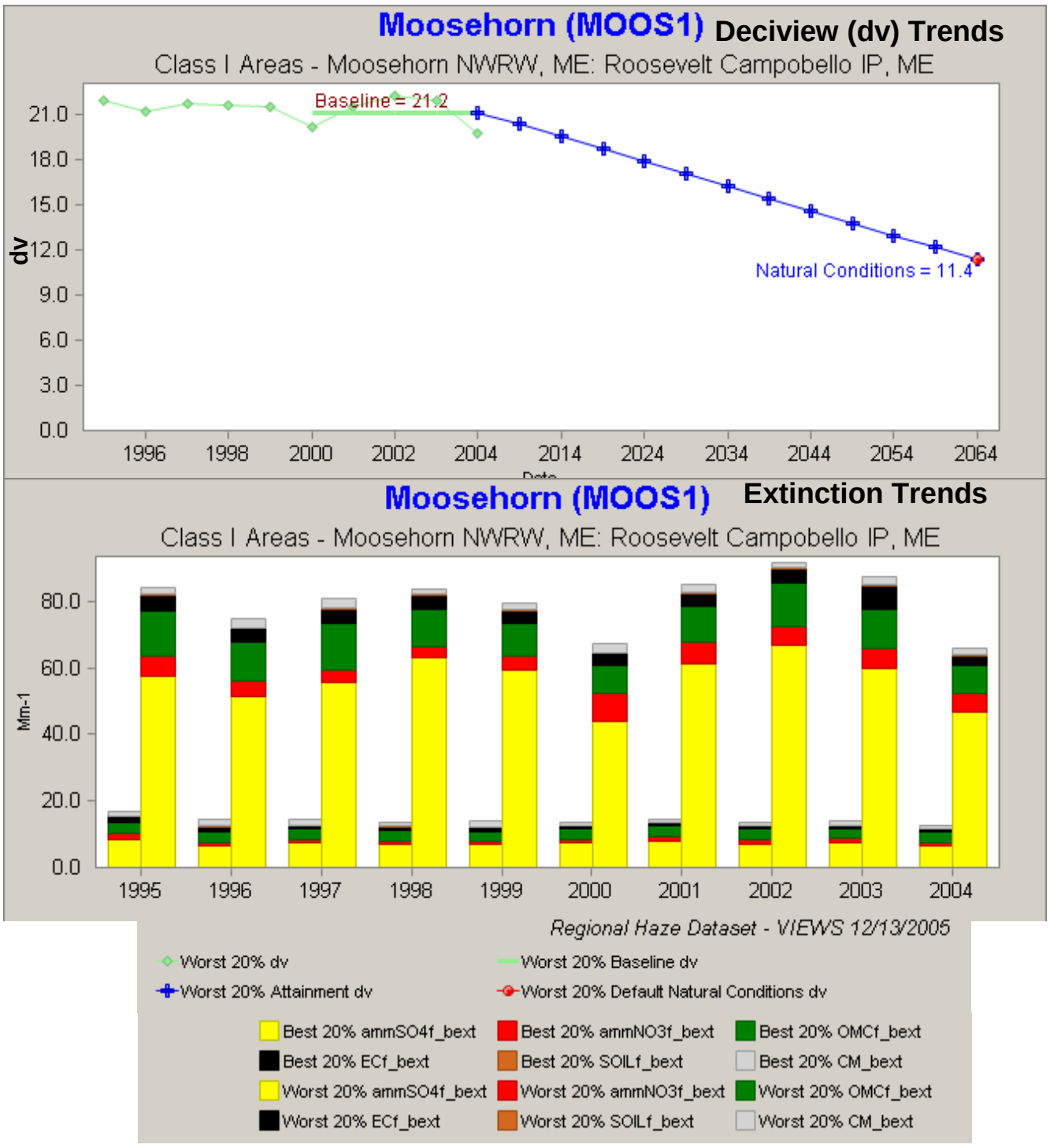


Figure B-6. Monitoring Data from Washington, DC

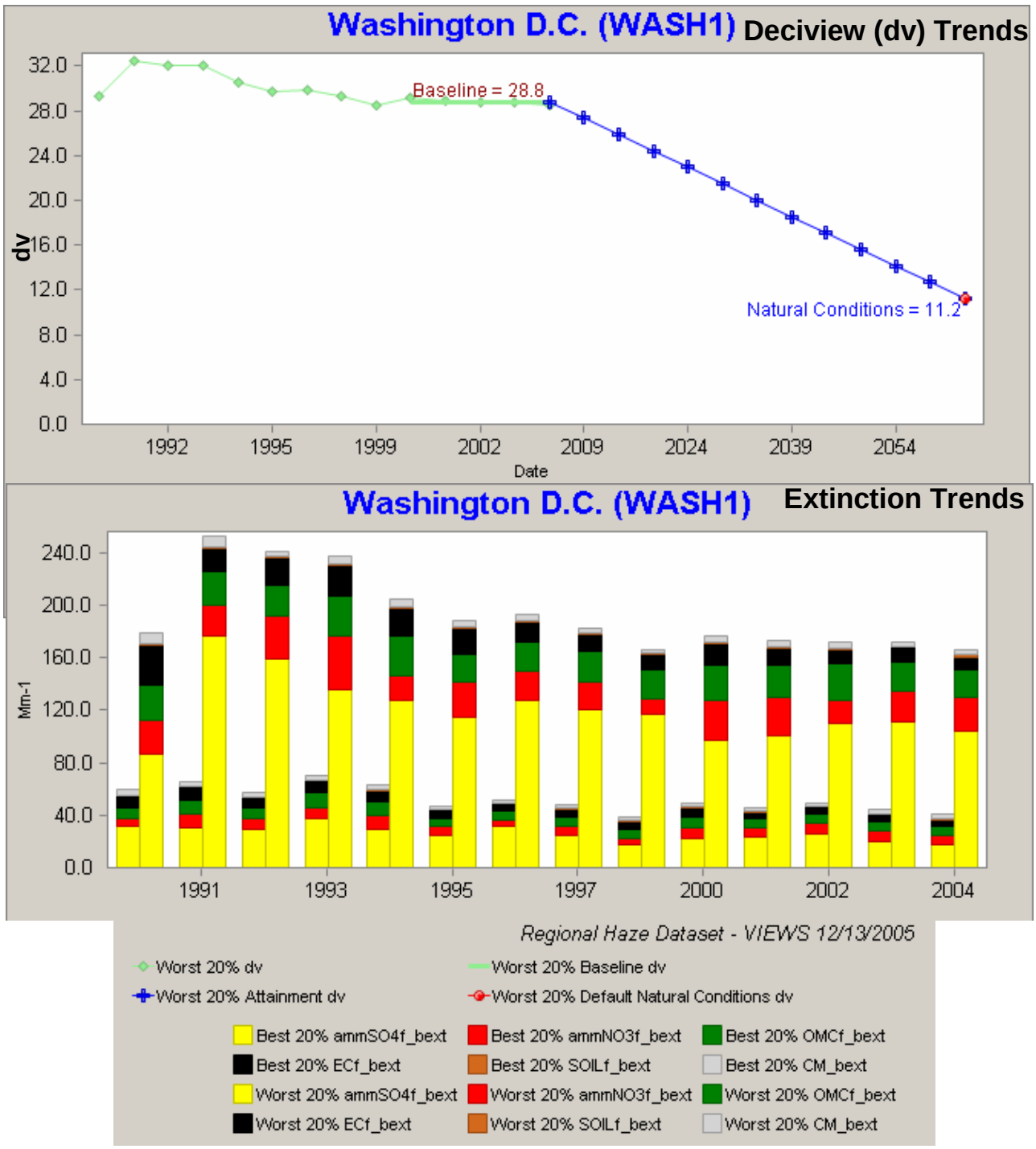
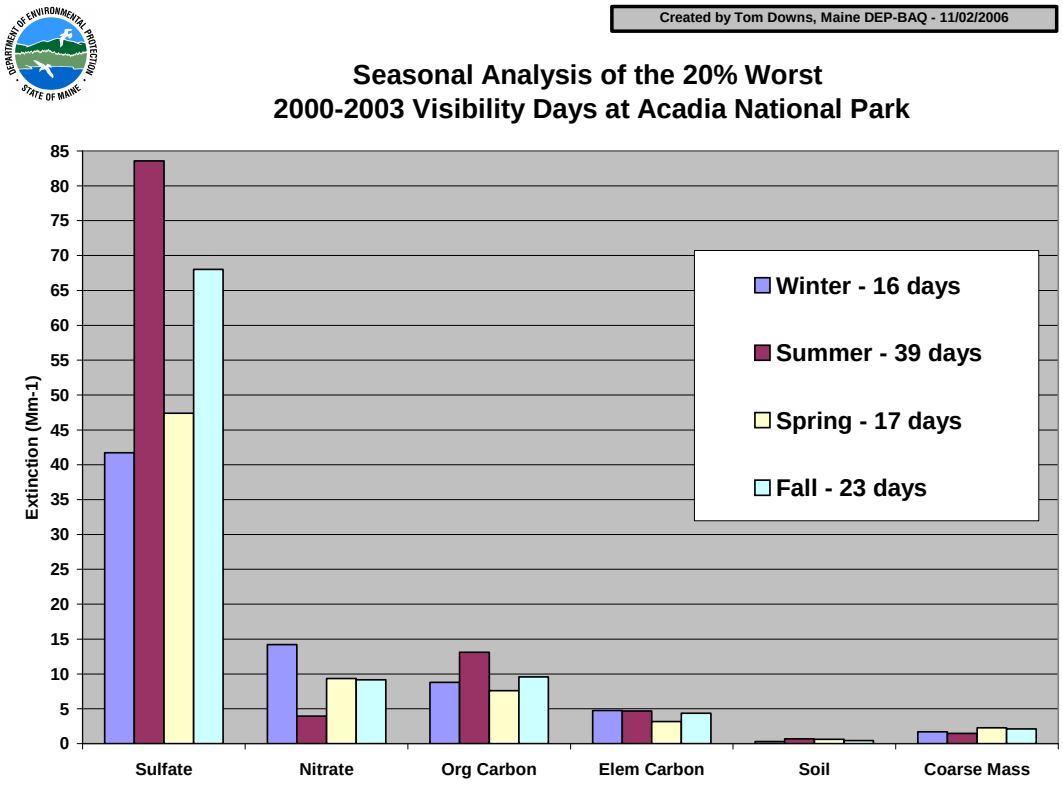


Figure B-7. 20% Worst and Best 2000-2003 Visibility Days at Acadia NP, ME



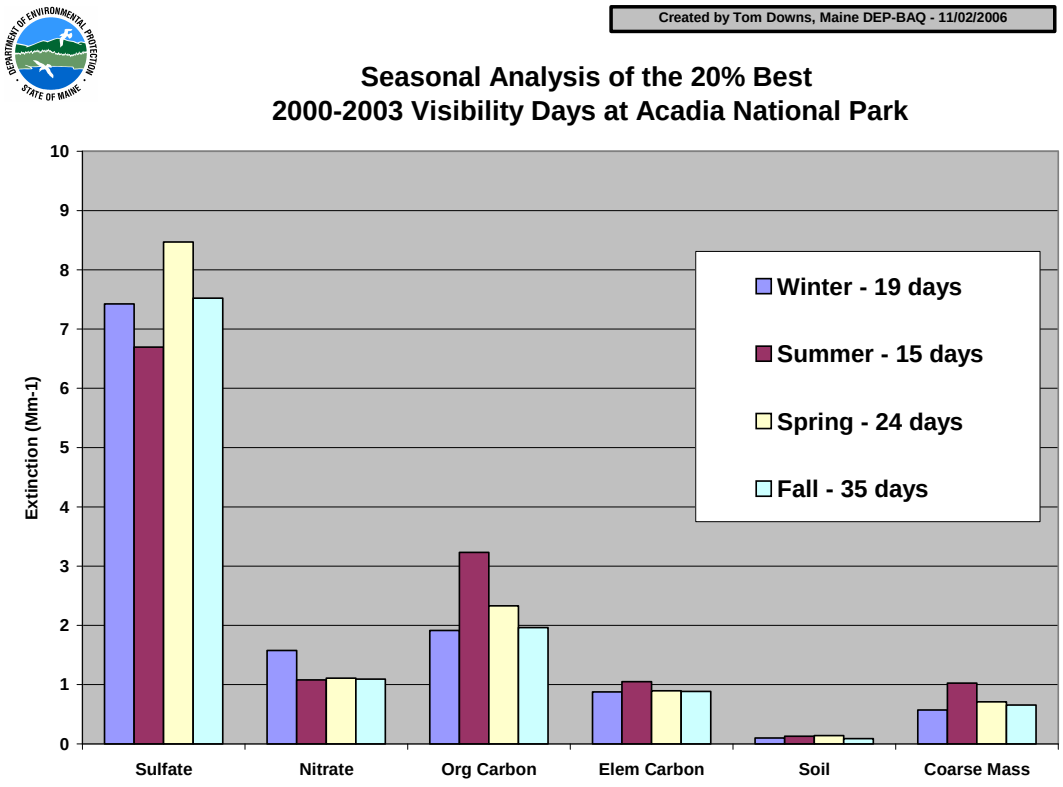
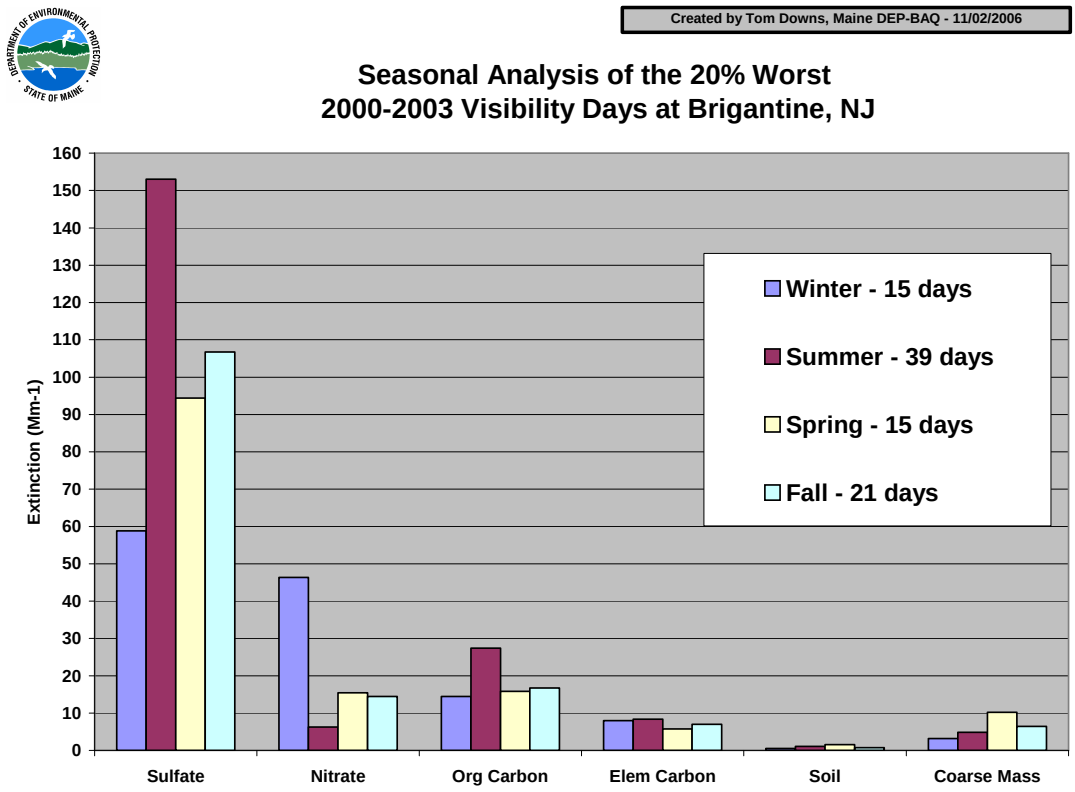


Figure B-8. 20% Worst and Best 2000-2003 Visibility Days at Brigantine, NJ



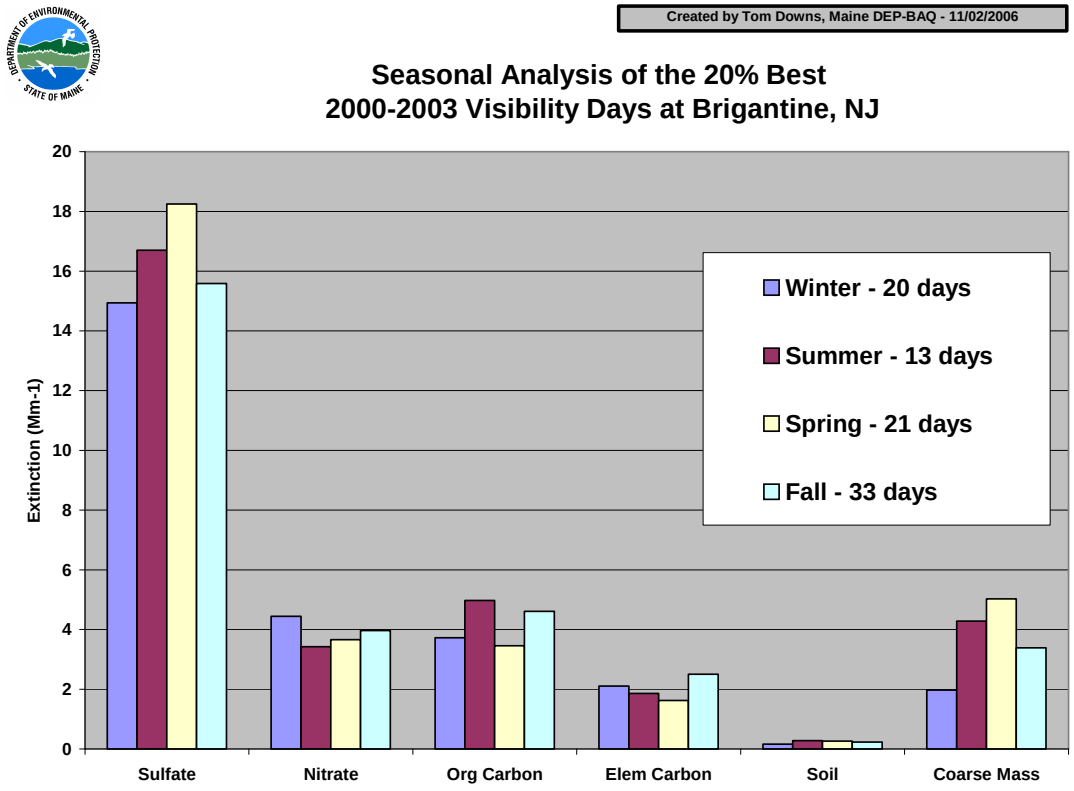
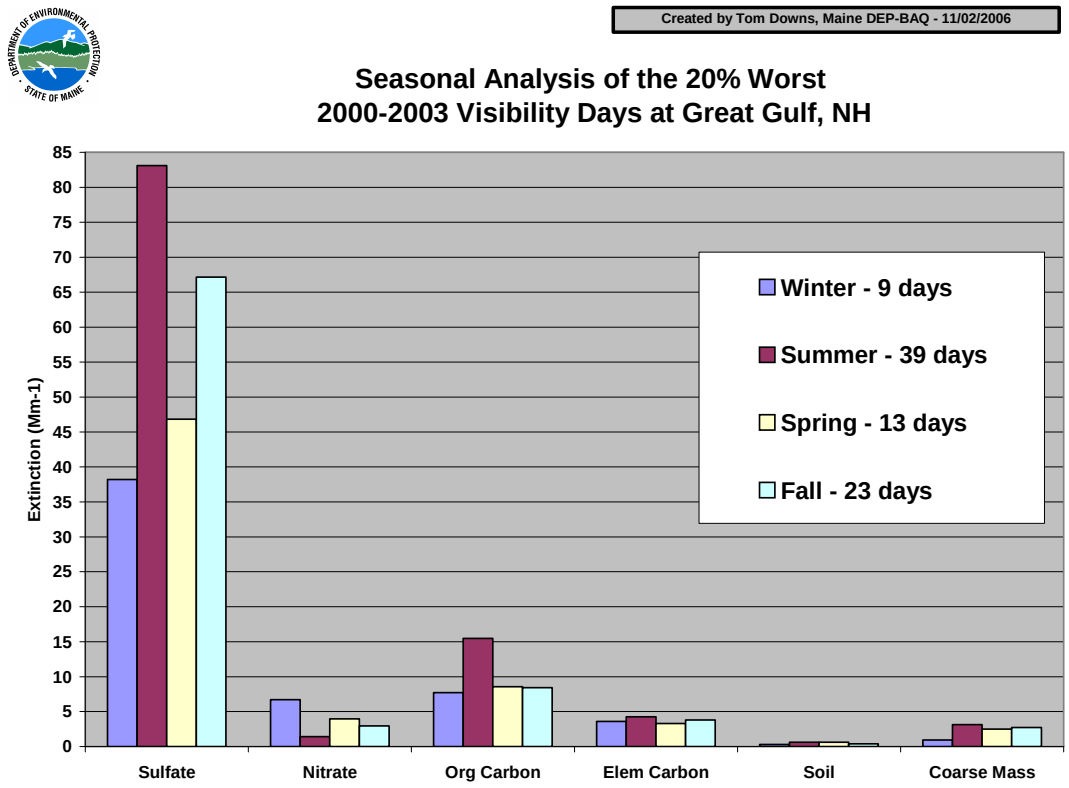


Figure B-9. 20% Worst and Best 2000-2003 Visibility Days at Great Gulf, NH



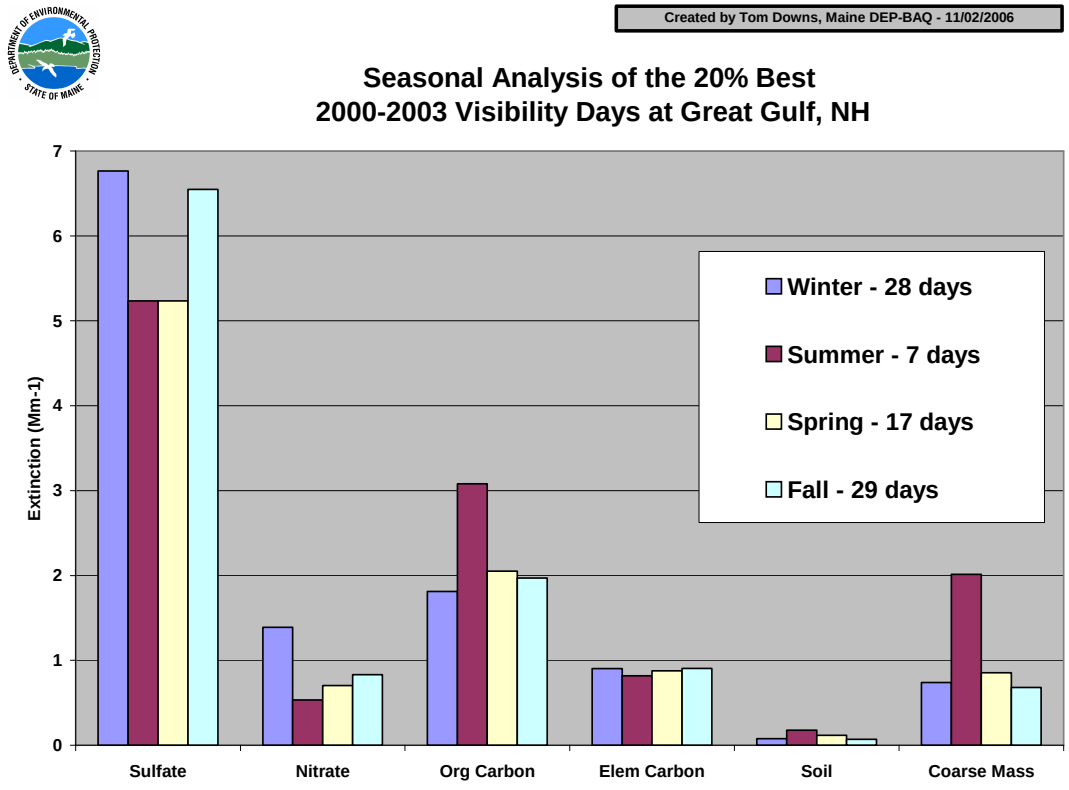
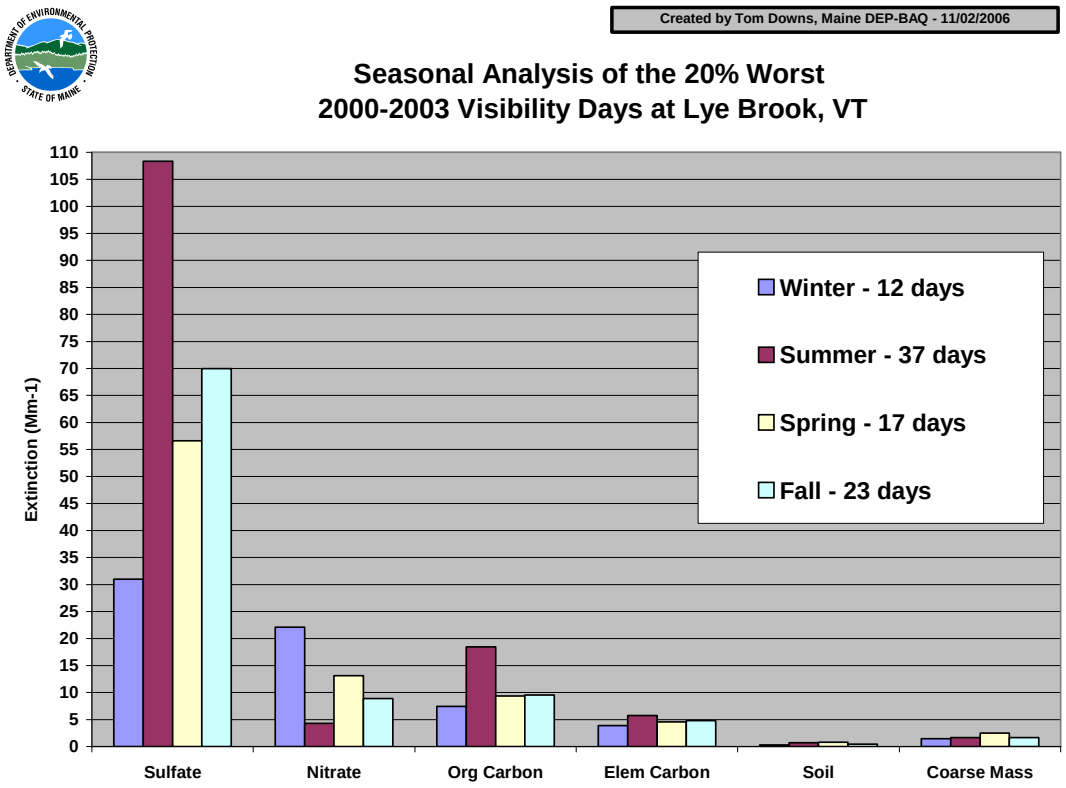


Figure B-10. 20% Worst and Best 2000-2003 Visibility Days at Lye Brook, VT



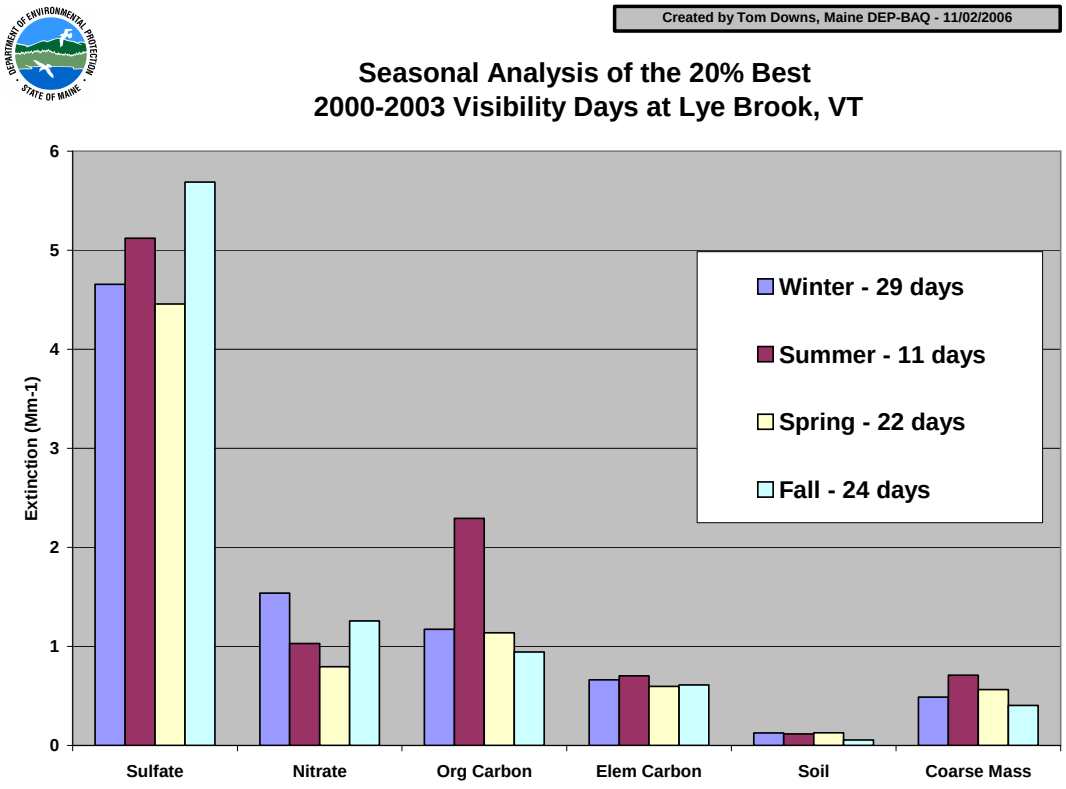
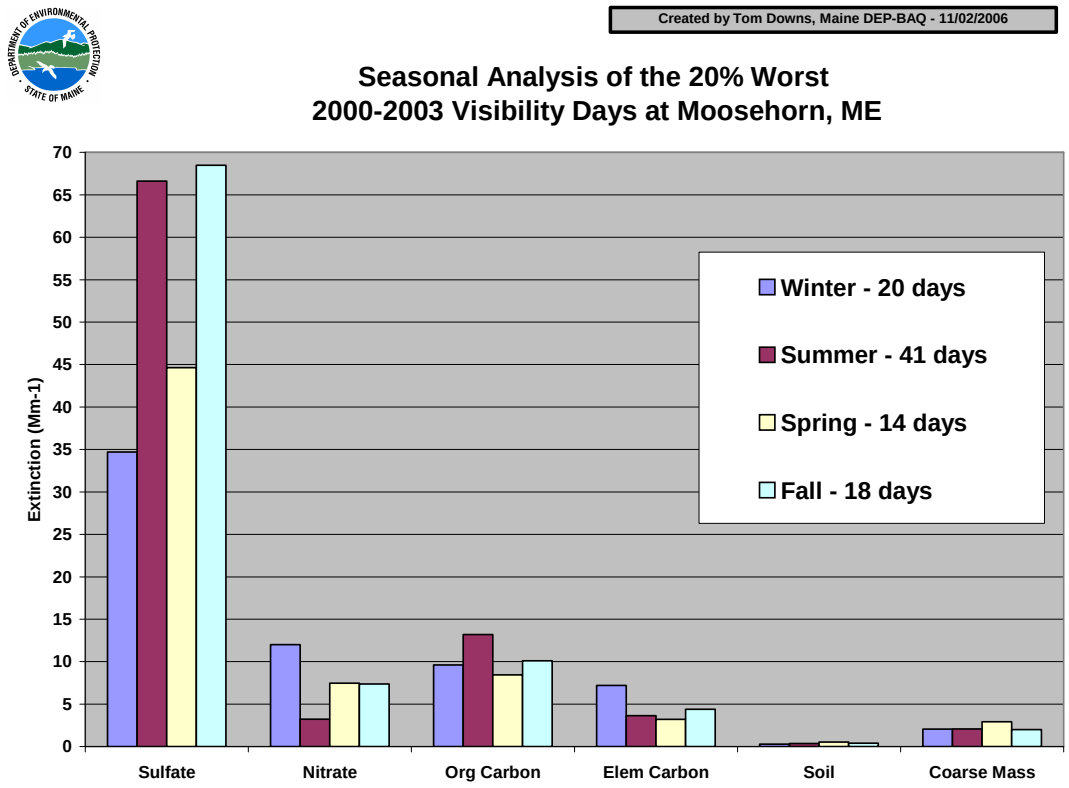


Figure B-11. 20% Worst and Best 2000-2003 Visibility Days at Moosehorn, ME



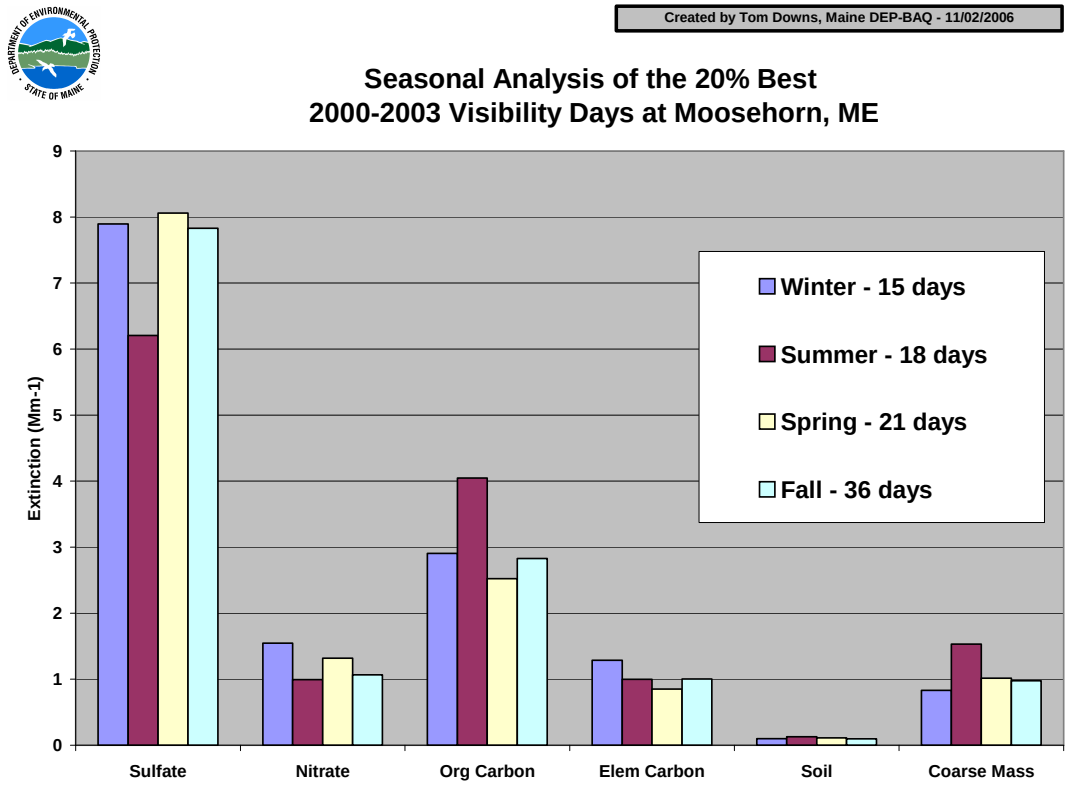
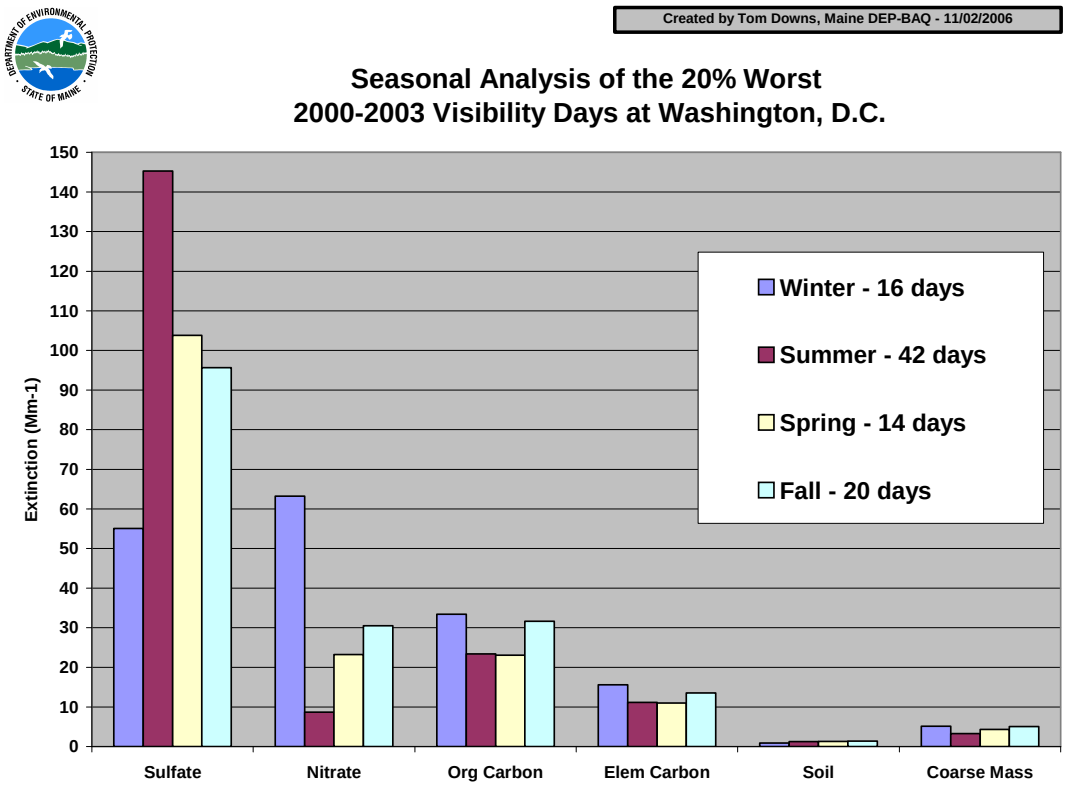


Figure B-12. 20% Worst and Best 2000-2003 Visibility Days at Washington, D.C.



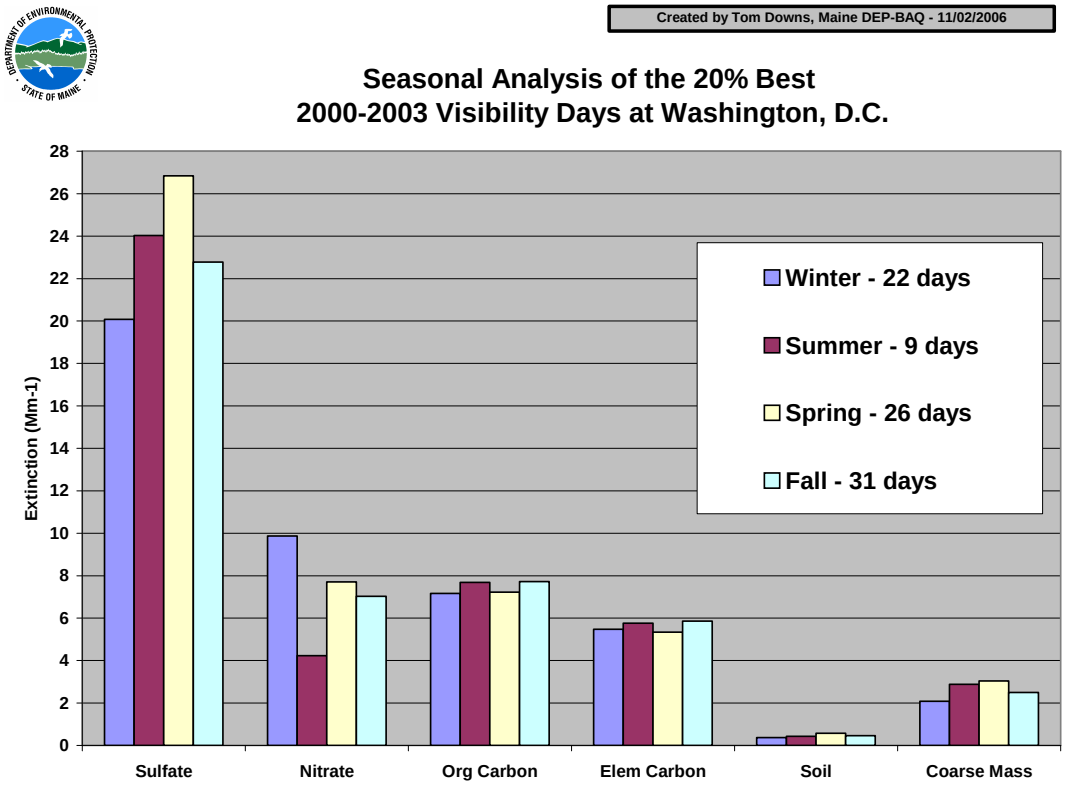
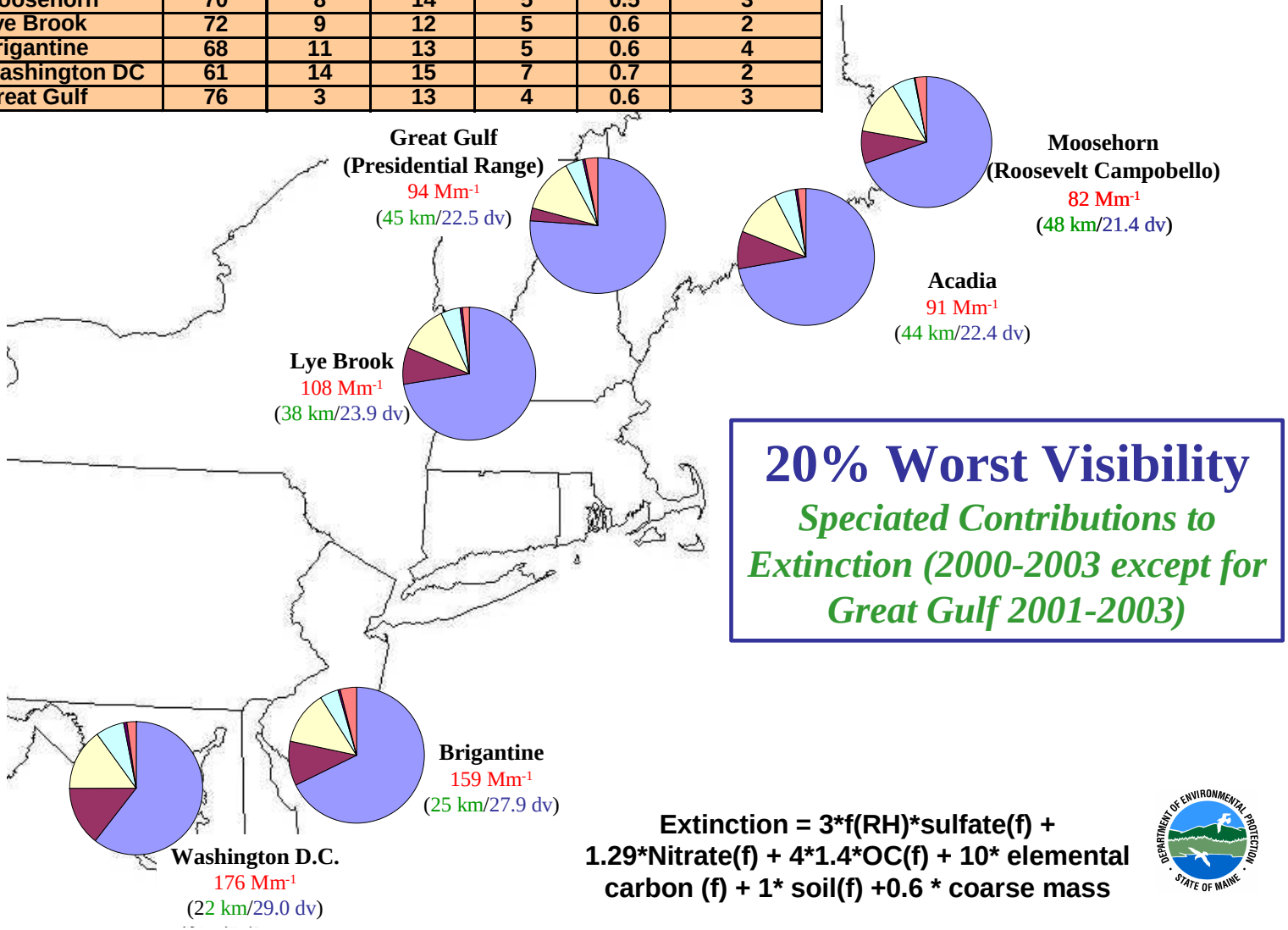


Figure B-13. 20% Best 2000-2003 Visibility Days Speciated Contributions to Extinction

	percent contributon to particle extinction					
Site	Sulfate	Nitrate	Org C	Elem C	Soil	Coarse Mass
Acadia	72	9	11	5	0.6	2
Moosehorn	70	8	14	5	0.5	3
Lye Brook	72	9	12	5	0.6	2
Brigantine	68	11	13	5	0.6	4
Washington DC	61	14	15	7	0.7	2
Great Gulf	76	3	13	4	0.6	3

Created by Tom Downs, Maine DEP-BAQ 12/13/2005



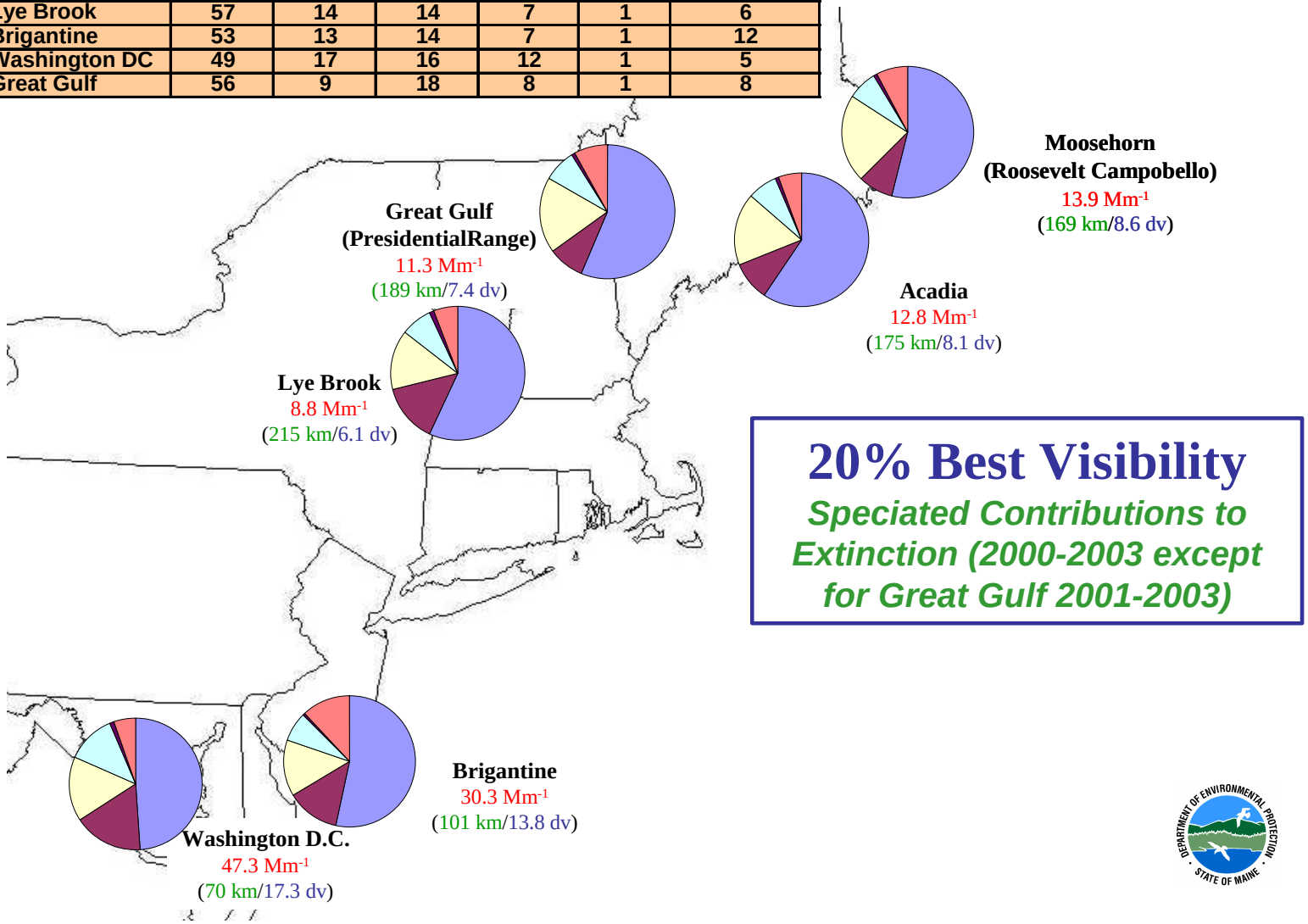
20% Worst Visibility

Speciated Contributions to Extinction (2000-2003 except for Great Gulf 2001-2003)

Extinction = 3*f(RH)*sulfate(f) + 1.29*Nitrate(f) + 4*1.4*OC(f) + 10* elemental carbon (f) + 1* soil(f) +0.6 * coarse mass

Figure B-14. 20% Best 2000-2003 Visibility Days Speciated Contributions to Extinction

Site	percent contributon to particle extinction					
	Sulfate	Nitrate	Org C	Elem C	Soil	Coarse Mass
Acadia	60	9	18	7	1	6
Moosehorn	54	9	22	7	1	8
Lye Brook	57	14	14	7	1	6
Brigantine	53	13	14	7	1	12
Washington DC	49	17	16	12	1	5
Great Gulf	56	9	18	8	1	8



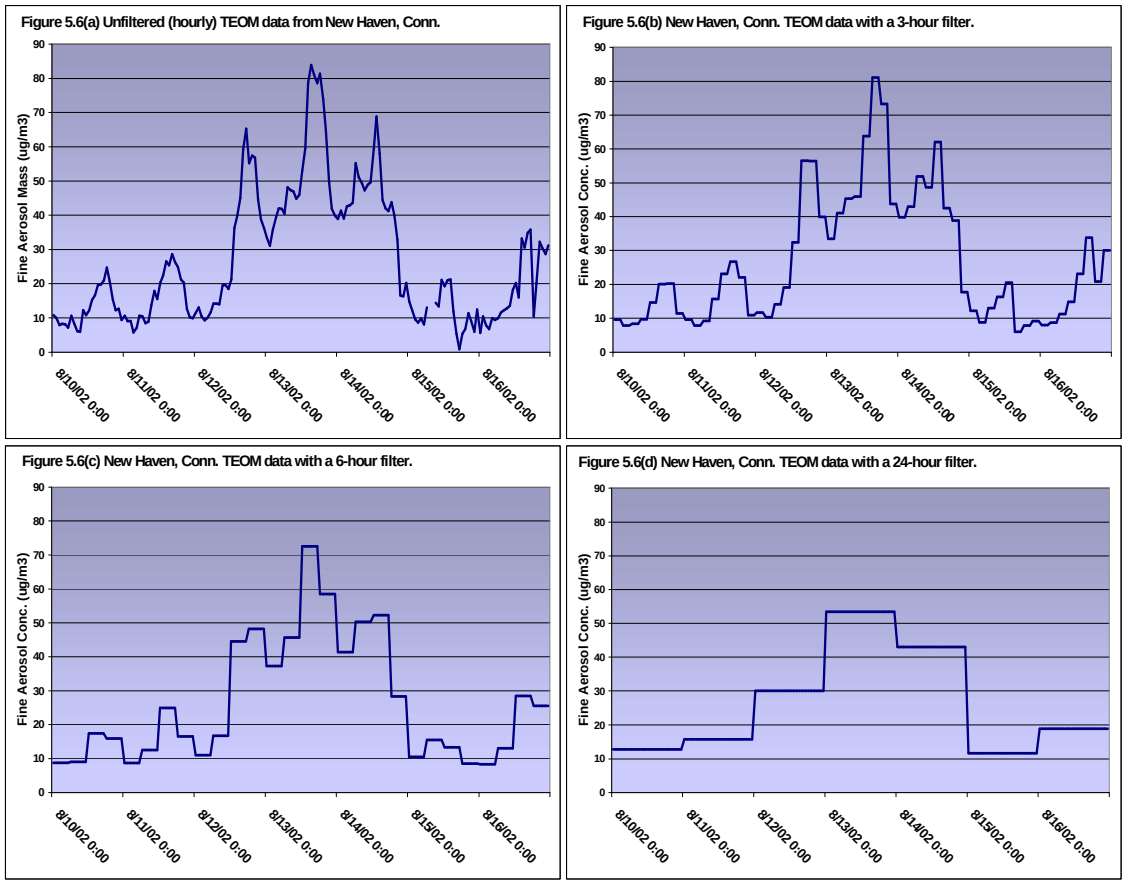
Appendix C: Additional Considerations for $PM_{2.5}$ Air Quality Management

APPENDIX C: ADDITIONAL CONSIDERATIONS FOR $PM_{2.5}$ AIR QUALITY MANAGEMENT

C.1. Averaging times and data interpretation

In analyzing the chemical data available for interpreting the air quality event of August 2002, it is important to point out that the use of different averaging times can have a profound effect on our understanding of the progression of any specific episode. Many subtleties of synoptic-scale meteorology and atmospheric chemistry are "aliased out" of data sets with temporal resolution greater than 3-6 hours. These effects are demonstrated in Figure C-1 which show fine aerosol TEOM data from New Haven for the "episode" period August 10-16, 2002. In these figures, the hourly TEOM values have been aggregated into 3-, 6- and 24-hour mean values. Average concentrations are inversely proportional to the length of the averaging period and the ratio of peak hourly concentration within a daily average ranges from about 1.5 to 1.75 for this episode.

Figure C-1. Effects of averaging times (or temporal resolution) on time series information



C.2. Rural versus urban $PM_{2.5}$ mass

Comparison of $PM_{2.5}$ concentrations from rural areas with those from urban/suburban areas can add significantly to our understanding of the impact on air quality of both urban sources and of medium to long-range fine aerosol transport. To assist with this approach, data from 10 pairs of rural and urban/suburban FRM sites throughout the MANE-VU region were selected and analyzed.

Table C-1 shows basic site description information including the approximate, straight-line distance between the site pairs.

Due to the difficulty in finding a significant number of urban-rural site pairs that operated on the same sampling schedule, sites with a mixture of schedules were used to insure samples representative of the entire MANE-VU region. As a result, three of the 20 sites employed an everyday schedule while two sites sampled every sixth day (the remainder sampled every third day). Data from the three everyday sites were edited so as to include data from the 1-in-3 schedule only. In all, a total of 1098 data points were possible from the 10 site pairs for 2002. Of the 1098 possible point-pairs, 951 (87%) were valid and were used in this analysis.

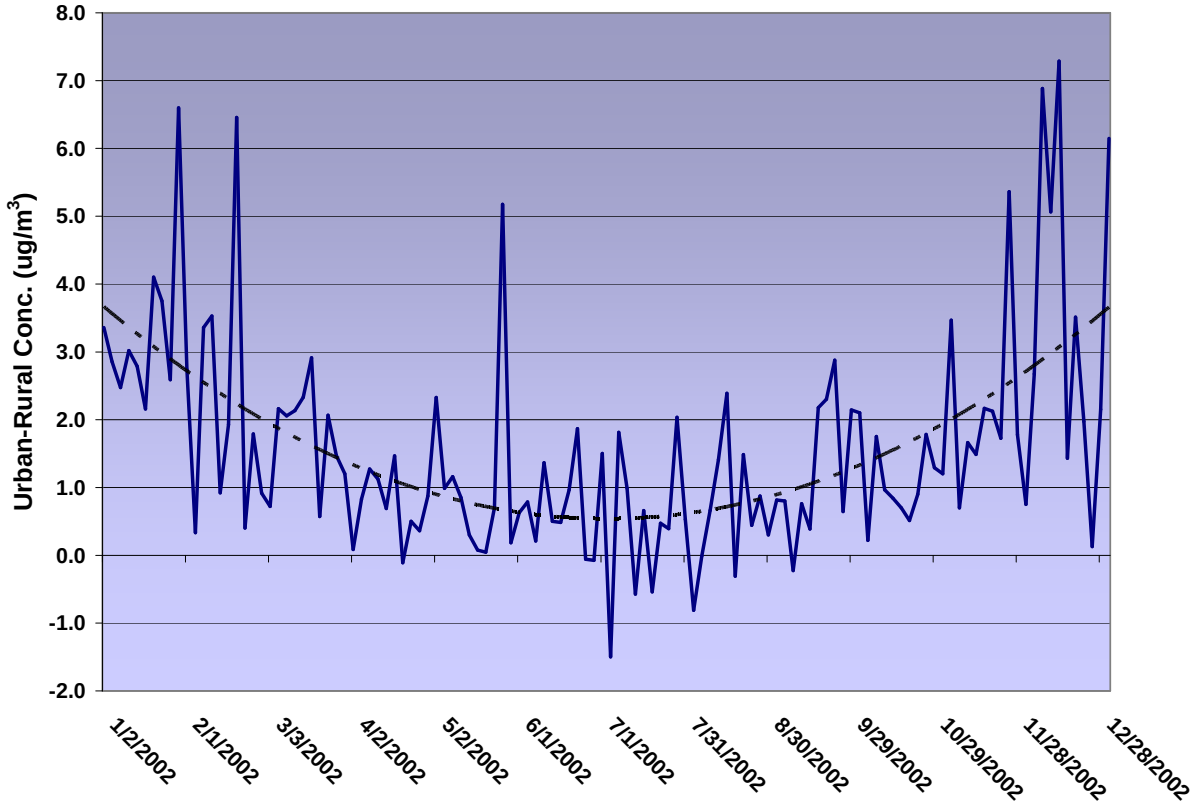
Table C-1. MANE-VU urban-rural site pair information

State	Site No	City	Land use	Location type	Longitude	Latitude	Inter-site Distance (mi)
DE	100051002		Agricultural	Rural	-75.55560	38.98470	
DE	100010002	Seaford	Residential	Suburban	-75.61310	38.64440	24.0
MA	250154002	Ware	Forest	Rural	-72.33472	42.29833	
MA	250130016	Springfield	Commercial	Urban & Center City	-72.59140	42.10890	17.6
MD	240030014		Agricultural	Rural	-76.65310	38.90250	
MD	245100049	Baltimore	Residential	Urban & Center City	-76.63750	39.26170	25.2
ME	230052003	Cape Elizabeth	Residential	Rural	-70.20778	43.56083	
ME	230010011	Lewiston	Commercial	Urban & Center City	-70.21500	44.08940	37.0
NJ	340218001		Agricultural	Rural	-74.85470	40.31500	
NJ	340210008	Trenton	Residential	Urban & Center City	-74.76360	40.22220	7.7
NY	360010012	Albany	Agricultural	Rural	-73.75690	42.68070	
NY	360930003	Schenectady	Residential	Suburban	-73.94020	42.79960	11.7
NY	361030001	Babylon	Commercial	Rural	-73.42030	40.74580	
NY	360590013	Bethpage	Residential	Suburban	-73.49060	40.76080	3.3
NY	360130011	Westfield	Agricultural	Rural	-79.60250	42.29080	
PA	420490003	Erie	Commercial	Suburban	-80.03860	42.14180	22.2
PA	420030093		Residential	Rural	-80.02080	40.60720	
PA	420030021	Pittsburgh	Residential	Suburban	-79.94140	40.41360	14.0
PA	420290100		Commercial	Rural	-75.76860	39.83440	
DE	100031012	Newark	Residential	Suburban	-75.76170	39.69190	10.0

As expected, urban/suburban areas, with their rich supply of emission sources, almost always reported higher concentrations than their nearby sister sites in rural areas. Of the 951 valid data pairs, 660 showed higher urban/suburban levels while 291 cases showed higher rural levels.

One interesting aspect of the 2002 urban-rural data concerns the pattern in seasonal differences between such site pairs. Figure C-2 shows the difference (urban-rural) between the 10 site pairs as a time series.

Figure C-2. Difference in FRM data between 10 urban-rural site pairs for 2002



Although some rural-to-urban seasonal differences are to be expected, the variation in the magnitude of this difference is surprising. In the warm/hot months, the mean rural/urban difference amounts to no more than ~0.7 µg/m^3 (based on a best-fit 2nd order polynomial curve), which is a relatively small differential. However, during the cool/cold months that difference climbs to almost 4 µg/m^3, demonstrating a total annual seasonal variation of at least 3 µg/m^3. Because the mean annual concentration of all sites is 12.6 µg/m^3, an annual variation of 3 µg/m^3 becomes significant.

One explanation for the observed seasonal variation concerns the temporal distribution of local and transported emissions. In the summertime, MANE-VU sites repeatedly experience sulfate events due to transport from regions to the south and west. During such events, rural and urban sites throughout MANE-VU record high (i.e., >15 µg/m^3) daily average $PM_{2.5}$ concentrations. During summer stagnation events, atmospheric ventilation is poor and local emissions are added to the transported burden with the result that concentrations throughout the region (rural and urban) are relatively

uniform. There are enough of these events to drive the urban-rural difference down to less than 1 $\mu g/m^3$ during warm/hot months.

During the wintertime, strong local inversions frequently trap local emissions during the overnight and early morning periods, resulting in elevated urban concentrations. Rural areas experience those same inversions but have relatively fewer local sources so that wintertime concentrations in rural locations tend to be lower than those in nearby urban areas. Medium and long-range fine aerosol transport events do occur during the winter but at a much reduced rate compared to summertime. So, it is the interplay between local and distant sources as well as meteorological conditions that drive the observed seasonal urban-rural difference in FRM concentrations.

C.3. Seasonal relationship between $PM_{2.5}$ and NO_X

Because nitrogen oxides (NO_X) can be a good indicator of regional as well as local emissions, NO_X data for the MANE-VU region was downloaded from USEPA's AQS. Ultimately, data from six widely separated MANE-VU NO_X sites were selected (one site each in CT, DC, MA, NH, PA and VT). Sites were selected both for high data capture rates and geographic location. The NO_X data were then aggregated into regional averages on a daily basis and compared to $PM_{2.5}$ FRM data from 34 "everyday" sampling sites (which were also averaged on a regional basis).

During 2002, there were virtually no periods when regional mean $PM_{2.5}$ concentrations rose above 20 $\mu g/m^3$ and were not accompanied by rising (or already high) NO_X concentrations. However, as seen in Figure C-3, NO_X concentrations vary widely on an annual basis and tend to occur out-of-sync with fine particle concentrations.

Although the min/max extremes of these two pollutants are offset in time, they are highly correlated during some parts of the year. For example, Figure C-4 shows the regional $PM_{2.5}$ and NO_X data for the coldest (Jan., Feb., Nov., and Dec.) and hottest (May, June, July and Aug.) seasons of 2002. Wintertime NO_X and $PM_{2.5}$ concentrations are rather well correlated (r^2=0.67) while summertime concentrations are not at all linked. This dichotomy can be explained by several coincident effects including: 1) reduced UV radiation during cold months (which prevents photolysis of NO_2 to O_3); 2) the increase in space heating requirements from stationary sources (which preferentially increases morning NO_X emissions; increased NO_X emissions due to "cold-start" mobile source engines and 3) decreased mixing height depths due to reduced solar input (which allows morning concentrations to build quickly). Note that the Spring/Fall $PM_{2.5}$ vs. NO_X correlation (not shown) lies about mid-way between the winter/summer values shown in Figure C-4.

Figure C-3. Regional $PM_{2.5}$ and NO_X in 2002

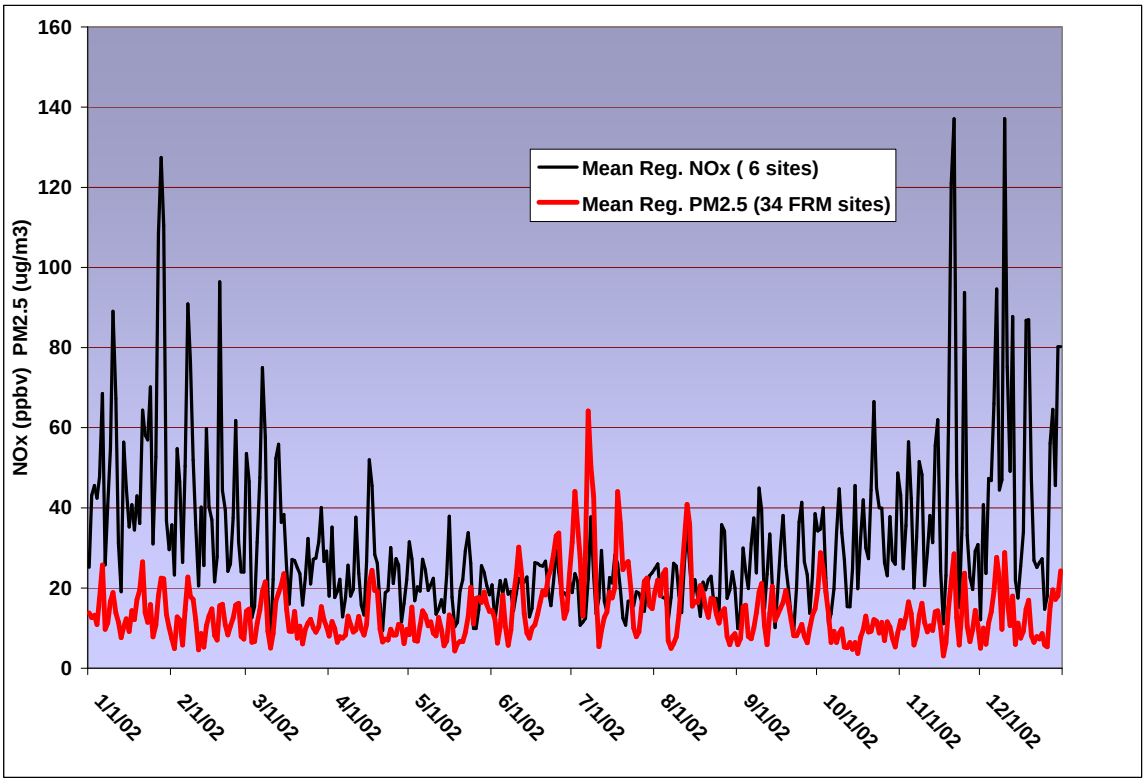
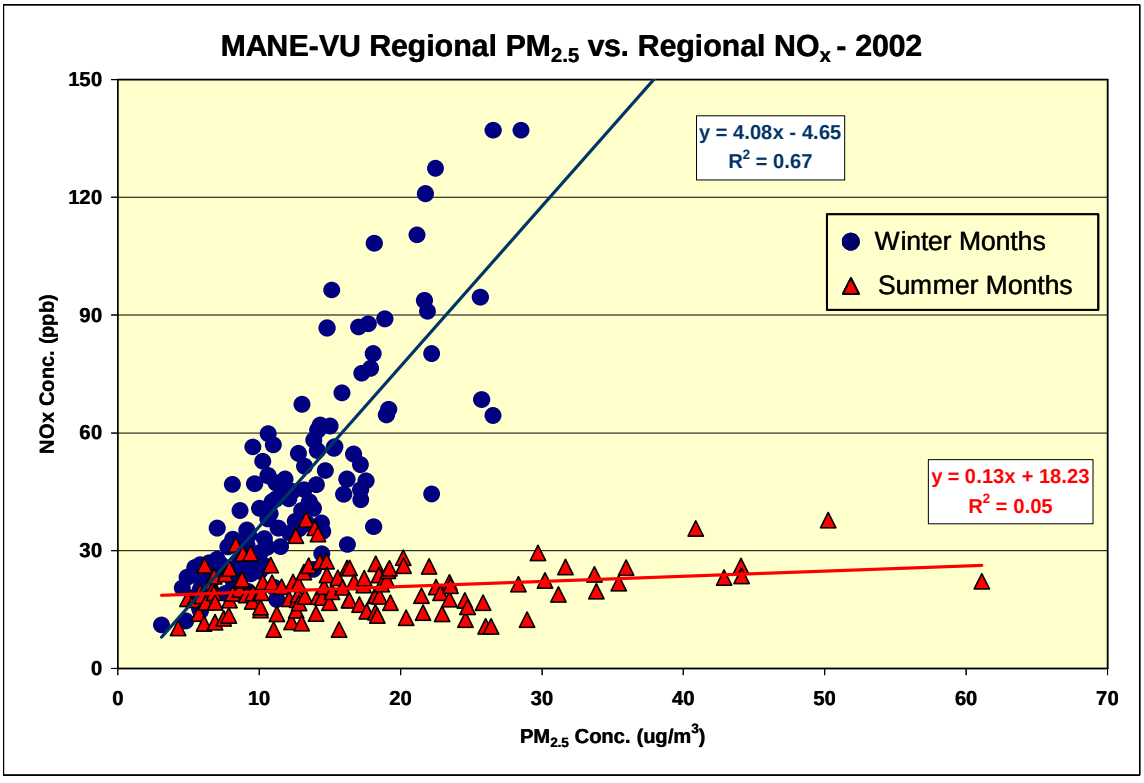


Figure C-4. $PM_{2.5}$ vs. NO_X correlation by season

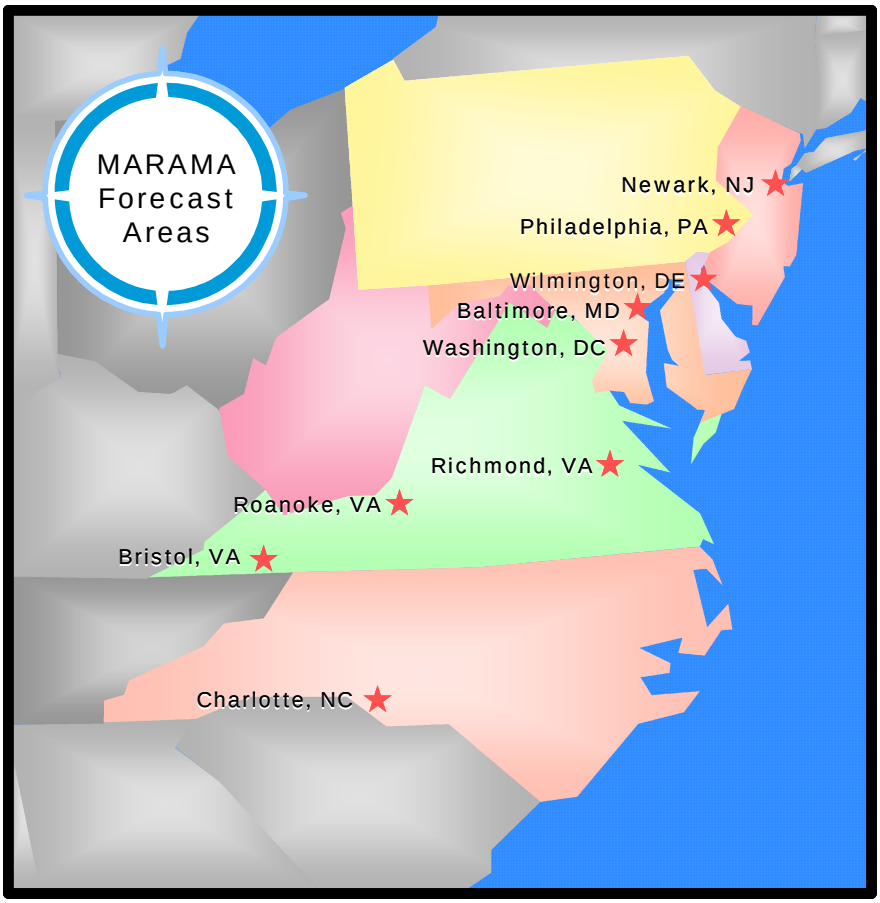


Mid-Atlantic Regional Air Management Association

The Development of $PM_{2.5}$ Forecasting Tools for Selected Cities in the MARAMA Region

Final Report

September 30, 2004



Prepared by
ICF Consulting/SAI
101 Lucas Valley Road, Suite 260
San Rafael, CA 94093
(415) 507-7108

Prepared for the
Mid-Atlantic Regional Air Management Association

04-046

About MARAMA

The Mid-Atlantic Regional Air Management Association is a voluntary, non-profit association of ten state and local air pollution control agencies. MARAMA's mission is to strengthen the skills and capabilities of member agencies and to help them work together to prevent and reduce air pollution impacts in the Mid-Atlantic Region.

MARAMA provides cost-effective approaches to regional collaboration by pooling resources to develop and analyze data, share ideas, and train staff to implement common requirements.

The following State and Local governments are MARAMA members: Delaware, the District of Columbia, Maryland, New Jersey, North Carolina, Pennsylvania, Virginia, West Virginia, Philadelphia, and Allegheny County, Pennsylvania.

About Systems Applications International, LLC

Systems Applications International (SAI), a division of ICF Consulting, Inc., specializes in air quality data analysis and modeling. SAI was established in 1968 and throughout its history has pioneered the use of advanced analytical and modeling tools to support air quality analysis and modeling assessments of primary and secondary pollutants.

SAI is the principal developer of the Urban Airshed Model modeling systems (UAM and UAM-V) and the REgional Modeling System for Aerosols and Deposition (REMSAD). SAI has also developed statistical data analysis techniques to support the selection and characterization of modeling episode periods and ozone and particulate matter forecasting.

On the Cover: Map showing the forecast areas of the nine $PM_{2.5}$ forecasting tools that were developed.

For copies of this report contact:

MARAMA
Mid-Atlantic Regional Air Management Association
711 West 40th Street
Suite 312
Baltimore, MD 21211

phone 410.467.0170
fax 410.467.1737
http://www.marama.org/

The Development of $PM_{2.5}$ Forecasting Tools for Selected Cities in the MARAMA Region

September 30, 2004

Prepared by
ICF Consulting/SAI
101 Lucas Valley Road, Suite 260
San Rafael, CA 94093
(415) 507-7108

Prepared for the
Mid-Atlantic Regional Air Management Association

04-046

Acknowledgements

This report was written by Sharon Douglas, Johanna Mangahas, Belle Hudischewskyj, and Jay Haney of ICF/SAI. Ana Alvarez, Geoffrey Glass, and Tom Myers of ICF/SAI are acknowledged for their technical contributions to the CART analysis and the development of the forecasting tools. This report was expertly produced by Jim Phillips. Special thanks to Bill Gillespie of MARAMA and all of the state and local forecasters listed below who contributed to the design and evaluation of the study components.

Rob Altenburg, Pennsylvania Department of Environmental Protection

George Bridgers, North Carolina Department of Environment and Natural Resources

Scott Jackson, North Carolina Department of Environment and Natural Resources

Joe Martini, Delaware NREC, Air Quality Management Section

Andy Mikula, New Jersey Department of Environmental Protection

Sean Nolan, Pennsylvania Department of Environmental Protection

Bill Ryan, Penn State University

Dan Salkovitz, Virginia Department of Environmental Quality

Matt Seybold, Maryland Department of the Environment

Ram Tangirala Ph. D, District of Columbia, Department of Health

Mike Woodman, Maryland Department of the Environment

Table of Contents

List of Figures

List of Tables

Executive Summary

The primary objective of the MARAMA $PM_{2.5}$ forecasting assistance project was to develop and evaluate statistical-based tools to support $PM_{2.5}$ forecasting for nine cities in the MARAMA region. The nine cities included Charlotte, North Carolina; Bristol, Roanoke, and Richmond, Virginia; Washington, D.C.; Baltimore, Maryland; Philadelphia, Pennsylvania; Wilmington, Delaware; and Newark/Elizabeth, New Jersey. The study included the analysis of PM2.5 and meteorological data using Classification and Regression Tree (CART) analysis software and the development, testing, and evaluation of interactive forecasting tools for each area. Data and information gathered throughout the course of the project were used, together with the CART analysis results, to describe the relationships between meteorology and PM concentration and, specifically, the conditions associated with high $PM_{2.5}$ events in each forecast area.

CART Analysis

The CART analysis software was applied for each area for a multi-year period (nominally 1999–2002). All days with available data within this period were classified and grouped into bins in accordance with the values of observed and calculated meteorological and air quality parameters that comprised the input dataset. Twenty-four-hour average $PM_{2.5}$ concentration was used as the classification variable for this application and a variety of meteorological and air quality parameters were used as input data.

The air quality data used for this study consisted of measurements of $PM_{2.5}$ from sites located within and potentially upwind of each area of interest. The final dataset used for the CART analysis included Federal Reference Method (FRM) $PM_{2.5}$ data.

The meteorological data used for this study consisted of measurements of various surface and upper-air meteorological parameters for sites located within and near each area of interest.

Each CART classification bin was assigned to one of three classification categories, representing a different range of $PM_{2.5}$ concentration. The three categories were defined according to the EPA established guidelines for PM2.5 forecasting: less than 15.5 (Category 1), 15.5 to less than 40.5 (Category 2), and greater than or equal to 40.5 μgm^{-3} (Category 3). Since only a few data points were in the highest EPA category of greater than or equal to 65 μgm^{-3}, this category was not used in the analysis. The three categories used in this analysis are also referred to as “good”, “moderate”, and “unhealthy for sensitive groups (USG).”

As part of the CART application, more than 20 diagnostic and sensitivity tests were conducted for each area. The first of these included only the meteorological input parameters. The remaining tests examined the use of alternate input parameters, as well as different forms of the classification variable for $PM_{2.5}$.

Key findings from the CART analysis include:

- Different types of $PM_{2.5}$ episodes can be identified for each area based on meteorological and prior day PM indicators.
- Regional $PM_{2.5}$ parameters are more important in classifying the days for smaller/southern urban areas; local $PM_{2.5}$ variables are more important for the larger/more northern areas.
- Stability parameters are important for all areas and more stable conditions are generally associated with higher $PM_{2.5}$ concentrations.

- Temperature is used to segregate the days seasonally and is overall well correlated with the observed $PM_{2.5}$ concentrations.
- Relative humidity is also used to segregate the days but high relative humidity can be associated with both low and high observed $PM_{2.5}$ concentrations
- Wind speed is important in defining classification groupings and lower wind speeds almost always lead to higher $PM_{2.5}$ bins.
- Wind direction is often used by CART to separate and group the days, but does not always vary regularly among the categories.
- For all areas, less precipitation is associated with lower $PM_{2.5}$ but is not frequently used by CART.

The CART results can be characterized in terms of classification accuracy, which is used to quantify the degree to which days within each bin have observed concentrations corresponding to the range assigned to the bin. Misclassification can occur due to a number of reasons including: monitoring network limitations, length (completeness) of the analysis period, use of discrete classification categories, and data errors or missing data.

For this study, two sets of final CART results were produced. The first of these was used to prepare the operational versions of the forecasting tools for each area. For this set of results, the average classification accuracy is 84 percent, ranging between 80 and 91 percent, as presented in Table ES-1.

Table ES-1. CART Classification Accuracy for the Operational Forecasting Tools

	Number of CART Bins	Classification Accuracy (%)
Charlotte	33	81
Bristol	33	90
Roanoke	34	91
Richmond	29	83
Washington	38	80
Baltimore	34	80
Philadelphia	35	82
Wilmington	36	81
Newark	34	86

For forecasting purposes, it is important that higher PM days are correctly classified, and that the number of lower PM days placed into higher PM bins is minimized. For Charlotte, Bristol, Roanoke, and Richmond, there were very few Category 3 days. All Category 3 days were correctly classified. There was some tendency for CART to place Category 1 and 2 days into the Category 3 bins, especially for Charlotte and Bristol.

For the Washington, Baltimore, Philadelphia, Wilmington, and Newark areas, there were more Category 3 days. With the exception of two days for Washington, all of the Category 3 days

were correctly classified. However, a significant number of Category 2 days (as well as some Category 1 days) were misclassified as Category 3.

A second set of CART results were produced for research purposes and were used to prepare research versions of the forecasting tools for each area. The research CART results differ from the operational CART results in their use of prior-day $PM_{2.5}$ input parameters. The research tools rely primarily on $PM_{2.5}$ data for one day rather than two days prior to the analysis day. For the research results, the average classification accuracy was 84 percent, ranging from 78 to 91 percent. Although the overall accuracy was similar, these results were generally less promising than the operational results, mostly because more days from Categories 1 and 2 were misclassified into the Category 3 bins. This is somewhat puzzling since more information about prior day $PM_{2.5}$ concentration should improve the classifications rather than degrade them. This issue was not resolved as part of the current project and the research versions of the tools were developed to allow further investigation of this issue and to support future work in this area.

$PM_{2.5}$ Forecasting Tools

The CART results were transformed into forecasting algorithms for each area so that observed and predicted values of the input parameters (for current and future days) could be used to place a future day into a classification bin. Future values for the meteorological parameters were obtained from standard meteorological forecast products, for example, the National Weather Service ETA model, the Global Forecast Systems (GFS) model, or the Nested Grid Model (NGM)). The resulting classification and forecast was determined by the observed and predicted data values and the pathways that comprise the CART classification tree. In this forecast mode, the predicted $PM_{2.5}$ concentration is assigned the value of the classification bin in which the day is placed.

This approach to forecasting has several attributes. Compared to simple regression techniques, the use a CART-based forecasting algorithm accommodates the possibility that different meteorological conditions can lead to the same or similar $PM_{2.5}$ concentration and, most importantly, that there may be multiple pathways to high $PM_{2.5}$. The parameters and parameter values associated with the CART classification tree provide information about the relative importance of these parameters in determining forecast $PM_{2.5}$ concentrations. Thus the CART technique offers physical insight into phenomena being studied. By segregating the data values into classification bins, CART also provides information regarding the frequency of occurrence of the conditions associated with each classification category. In this manner, the likely recurrence rate for a particular type of day and the associated prevalent conditions were obtained.

An important consideration in forecasting is, of course, the availability of real-time data to support forecasting. $PM_{2.5}$ data collected using FRM measurements were used for the CART analysis—as they were expected to provide the most consistent and accurate values. However, forecasters must rely on continuous measurements of $PM_{2.5}$ (which are available on a near real time basis) to provide information about prior day PM levels at local and upwind sites. Continuous $PM_{2.5}$ measurements do not always agree with the FRM measurements. Adjusting the continuous data to an FRM equivalent value may be one way to overcome this limitation.

For each set of CART results, four tools were developed. The four tools were for: 1) Charlotte; 2) Bristol, Roanoke, and Richmond; 3) Baltimore, Washington, Philadelphia, and Wilmington; and 4) Newark. Each tool consisted of an interface for the entry of observed and forecasted data and other parameters, the forecasting algorithms and supporting calculations for one or

more areas, and several options for the display, summary, and storage/archival of the input parameters and the forecast results. The operational versions of the tools were used to support the first year of $PM_{2.5}$ forecasting for several of the areas of interest.

Preliminary versions of each tool were evaluated on a real-time basis and using historical data. Meteorologists in six of the nine MARAMA areas tested the draft versions of the operational $PM_{2.5}$ forecasting tools during February and March of 2004. For as many days as possible, each participant entered the measured and forecasted meteorological and air quality data required by the tool to predict the next day's $PM_{2.5}$ level. For the real-time evaluation, prediction accuracy ranged from 55 to 75 percent using strict evaluation criteria, and from 75 to 88 percent when days with observed concentrations very close to the values defining the different categories were considered to be correctly classified within either category. It is important to keep in mind in reviewing these percentages that all of the days forecast exhibited low (good) or moderate $PM_{2.5}$ levels. No high PM days were observed at the continuous monitors in February and March of 2004.

Continuous data were used to evaluate the forecasts during the initial real-time evaluation period. Later, the evaluation statistics were recalculated for four of the areas using FRM data. Forecast accuracy was better for Richmond and Wilmington, but worse for Charlotte and Baltimore when the FRM data were used in place of the continuous data for evaluation. The greatest differences in performance were for Wilmington (where the FRM concentrations tended to be lower than the continuous values) and Baltimore (where the FRM concentrations tended to be higher than the continuous values). Thus, uncertainty in the observed PM concentrations may affect the integrity of the real-time evaluation results.

Overall, the real-time testing of the draft version of the forecasting tools was inconclusive primarily because the period February-March 2004 did not contain any days with high $PM_{2.5}$ concentrations.

Use of historical data for June through August 2003 enabled evaluation of the forecasting tools for all nine areas. Unlike the real-time 2004 evaluation, the summer 2003 period provided ample USG days to test the tools' ability to accurately predict high PM.

The historical evaluation suggested that given perfect forecasts of the meteorological input parameters, $PM_{2.5}$ concentration ranges can be correctly predicted for 50 to 70 percent of the days and correctly predicted using the less strict evaluation criteria for 65 to 85 percent of the days (with the exception of Bristol, which has a 55 percent accuracy even with the less strict criteria).

In the historical evaluation, the two sites with the worst performance were Bristol and Roanoke. These sites had fewer data than the other sites. This outcome suggests that, because of the limited database, the CART results are incomplete with respect to representing all of the types of conditions that might occur at these sites. The implication is that use of a limited dataset may limit the predictive capability of the tools. The limited size of the historical database used to develop the tools limits for forecast performance for all areas.

The false alarm rate was relatively high for all areas, where it could be calculated, and this reflects the tendency for overestimation found in the CART results. With this tendency, the probability of detection is good for most sites, and the bias is positive in all cases for which it could be calculated. This outcome suggests that the meteorological inputs, and consequently the CART results may not sufficiently represent the conditions associated with the day-to-day transition from high to lower PM concentrations. The overpopulation of the higher PM bins with

lower PM days (both in the CART results and in the historical forecast results) may also be due to a lack of a sufficient number of high PM days in the dataset. High $PM_{2.5}$ days are needed in the dataset to provided a good representation of the conditions that are associated with these days.

For a first attempt at developing a CART-based forecasting tool for these nine areas, the results are promising. The evaluation statistics are lower than, but not that much lower than those that would be considered good for 8-hour ozone forecasting (and ozone is a simpler pollutant to forecast and has been much more extensively measured and studied).

Factors Influencing $PM_{2.5}$ Concentrations

In describing the factors influencing $PM_{2.5}$ concentrations for each area, we considered 1) the magnitude and spatial and temporal characteristics of the $PM_{2.5}$ concentrations, 2) the meteorological features influencing $PM_{2.5}$ concentrations, and 3) the characteristics of high $PM_{2.5}$ events. The analysis was designed to complement, in a qualitative sense, the forecast information provided by the CART-based $PM_{2.5}$ forecasting tools.

With regard to the spatial and temporal characteristics of the $PM_{2.5}$ concentrations observed in the areas of interest:

- There is a greater incidence of high $PM_{2.5}$ days in the northern part of the MARAMA study area and within the larger metropolitan areas.
- During the period studied, the largest number of observed USG days occurred either during the second or third quarters of the year, encompassing the late spring and summer periods, although some USG days occurred during the fall and winter months as well in some areas.
- Correlations of $PM_{2.5}$ concentrations among the different areas suggest that there is a regional component to $PM_{2.5}$ in the areas of interest from Washington (possibly Richmond) northward, but that on any given day (with a few exceptions) there are also local meteorological and/or emissions influences that affect the areas separately.
- The characteristics of high $PM_{2.5}$ events vary among the areas of interest according to geographical characteristics and local and regional emissions characteristics.

Considering the meteorological features influencing $PM_{2.5}$ concentrations and the characteristics of high $PM_{2.5}$ events:

- A review of the meteorological conditions associated with high $PM_{2.5}$ in the areas of interest reveals that many of these days are influenced by a slow-moving or stationary high pressure system over the area of interest that results in suppressed vertical mixing of emissions/pollutants and low wind speeds or stagnation.
- For most of the areas, there are different types of high $PM_{2.5}$ events and these are distinguished by different stability characteristics and wind directions; the overall characteristics also vary with season.
- CART appears to be able to distinguish and group the USG days quite effectively.

Recommendations

Based on the results and findings of the study, as well as the issues and problems that we encountered in conducting the work, we provide recommendations for future enhancement of the forecasting tools and an improved understanding of $PM_{2.5}$ issues in this section.

All aspects of this study (including the development, refinement, and evaluation of the forecasting tools) emphasize the need for daily FRM and continuous $PM_{2.5}$ data on both a local and regional basis.

As more $PM_{2.5}$ data become available, use of a larger dataset encompassing a longer time period would likely better capture the range of different meteorological/$PM_{2.5}$ conditions that are likely to occur in the future as well as to better characterize the conditions associated with high PM days (which were few in number during the analysis period for several of the areas).

Continued evaluation of the forecasting tools, including an assessment of the different meteorological forecast products, will also provide important information related to improving forecast skill.

The data and results of this study could be used to enhance $PM_{2.5}$ State Implementation Plan (SIP) analyses for the areas of interest. Specifically this study can be used to support the development of a "conceptual description" of $PM_{2.5}$ formation and transport for each area (a required element of a SIP). The study results could also be used in "weight-of-evidence" analyses in which data and modeling results are use to support or corroborate the outcome of an attainment demonstration.

1. Introduction

1.1. Background and Objectives

The recent emphasis on fine particulate matter as an air pollutant of concern is based primarily on epidemiological studies that have indicated a cause and effect relationship between exposure to fine particles and health effects, including respiratory and cardiovascular disease and premature mortality. Particulates are also a primary constituent of regional haze, which limits visibility and thus diminishes the natural beauty of our environment.

Fine particulates in the atmosphere consist of primary particles that are emitted directly from sources and secondary particles that form in the atmosphere through chemical and physical processes. Pollutants that contribute to the formation of secondary aerosols include sulfur dioxide (SO_2), oxides of nitrogen (NO_x), and ammonia (NH_3). Natural sources of fine particulates and precursor pollutants include wind blown dust, sea salt, and forest fires. Anthropogenic contributors include numerous agricultural, mobile, and industrial sources. Meteorology plays an important role in particulate formation and transport and the determining the ambient particulate concentration levels.

The U.S. Environmental Protection Agency (EPA) established new National Ambient Air Quality Standards (NAAQS) for fine particulate matter in 1997. Under these standards, fine particles are defined as those with a diameter of less than 2.5 microns; particles of this size are also referred to as $PM_{2.5}$. The annual $PM_{2.5}$ standard requires the three-year average annual mean concentration to be less than 15 micrograms per cubic meter (μgm^{-3}). The daily $PM_{2.5}$ standard requires the three-year average of the 98^{th} percentile daily average concentration to be less than 65 μgm^{-3}. According to recent data and recommendations by the States and EPA, the Mid-Atlantic Regional Air Management Association (MARAMA) region contains several nonattainment areas for $PM_{2.5}$, based on the annual standard.

Compliance with these standards requires state and local agencies to monitor $PM_{2.5}$ concentrations within populated areas and, as needed, to develop and implement air quality plans for attainment and maintenance of the standards. To help protect public health, state and local agencies began daily forecasting of $PM_{2.5}$ concentrations in October 2003. Information regarding expected $PM_{2.5}$ concentrations allows the public to make informed decisions about their daily activities and to avoid unnecessary exposure to unhealthful concentrations. This information can also be used by businesses and industries to guide activities related to mitigation of emissions that may contribute to unhealthful particulate levels.

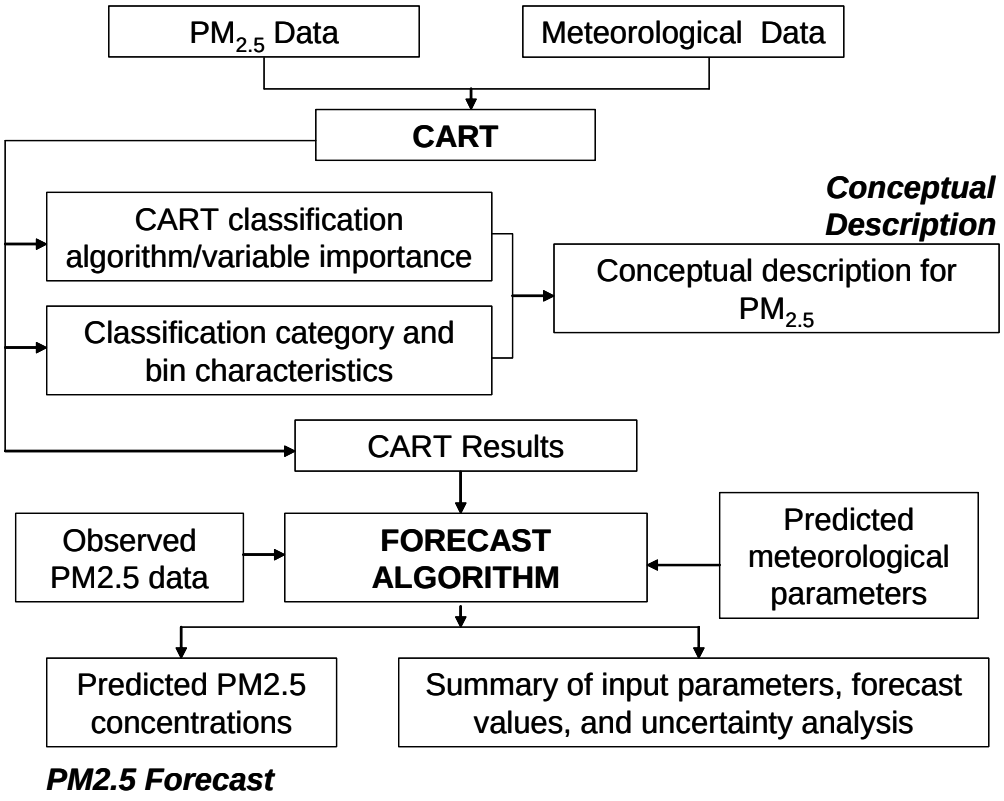
The primary objective of the MARAMA $PM_{2.5}$ forecasting assistance project was to develop and evaluate statistical-based tools to support $PM_{2.5}$ forecasting for nine cities in the MARAMA region. The nine cities include Charlotte, North Carolina; Bristol, Roanoke, and Richmond, Virginia; Washington, D.C.; Baltimore, Maryland; Philadelphia, Pennsylvania; Wilmington, Delaware; and Newark/Elizabeth, New Jersey. A secondary objective was to use available data and the results from the statistical analysis to understand the factors influencing $PM_{2.5}$ formation and transport in each forecast area and the MARAMA region.

1.2. Technical Overview of the Project

In this study, we used available air quality and meteorological data, together with the Classification and Regression Tree (CART) analysis technique, to develop forecasting

algorithms as well as a description of the factors influencing $PM_{2.5}$ concentrations within each of the nine areas of interest. The data were obtained from EPA and the National Weather Service (NWS) and were processed and quality assured for use in the CART analysis. The CART technique was then used to examine and extract information from the data, and the resulting information was used to describe each area and to develop the forecasting algorithms. A schematic diagram of the CART-based forecasting and analysis methodology is provided in Figure 1-1.

Figure 1-1. Conceptual Design of the MARAMA PM2.5 Forecasting and Analysis Methodology



CART is a statistical analysis tool that can be used to separate days with different values of a classification variable into different bins. The CART technique accomplishes this task through the growth of a binary decision tree, comprised of a progression of binary spits on the values of a set of input variables. The resulting tree has multiple branches, of varying complexity, each of which represents a path to a specific bin. Each bin is associated with a range of values of the classification variable.

For this analysis, CART was applied for a multi-year period (nominally 1999–2002) and all days within this period were classified and grouped into bins in accordance with the values of observed and calculated meteorological and air quality parameters that comprise the input dataset. We used 24-hour average $PM_{2.5}$ concentration as the classification variable for this application and a variety of meteorological and air quality parameters as input data.

The resulting CART trees were transformed into forecasting algorithms for each area so that observed and predicted values of the input parameters (for current and future days) can be used to place a future day into a classification bin. Future values for the meteorological parameters are obtained from standard meteorological forecast products. Using this approach, the path taken through the CART tree and the resulting classification is determined by the observed and predicted data values and the binary splits that comprise the classification tree. In this forecast mode, the predicted $PM_{2.5}$ concentration is assigned the value of the classification bin in which the day is placed. By providing a basis for estimating $PM_{2.5}$ concentrations using observed (or predicted) values of related variables, CART analysis can be used to forecast $PM_{2.5}$ concentrations.

The CART-based forecasting algorithm relies on the relationships that are identified between the input variables and $PM_{2.5}$ concentration (as derived using observed data). We also used this information in this study to improve our understanding of the factors and processes contributing to high $PM_{2.5}$ values in the areas of interest and throughout the region.

This approach enabled the preparation of useable forecasting tools to support the first year of $PM_{2.5}$ forecasting for several of the areas of interest. However, the ability of the tools to represent the type and range of conditions and the different types of $PM_{2.5}$ events that characterize each area is limited by the data used to develop the tool. Data for 1999–2002 were used, and, for most areas, data were available for only a subset of this period. It is anticipated that the incorporation of new data and information would enhance the performance of the tools as well as our understanding of $PM_{2.5}$ issues.

1.3. Report Contents

A summary of the data used in this project is provided in Section 2 of the report. The CART application is described in detail in Section 3. The factors influencing $PM_{2.5}$ concentrations within each area are discussed in Section 4. The forecasting tools are documented in Section 5, and an evaluation of the tools is presented. Finally, some recommendations for further study are provided in Section 6.

This page intentionally left blank.

2. Project Database

This project relied on historical air quality and meteorological data to support the development and evaluation of the CART-based forecasting tools. The acquisition, processing, and archival of the historical data is described in this section of the report.

2.1. Air Quality Data

The air quality data used for this study consist of measurements of $PM_{2.5}$ for sites located within and potentially upwind of each area of interest. The final dataset used for the CART analysis includes $PM_{2.5}$ data obtained using the Federal Reference Method (FRM) measurement systems. Data collected using one or more continuous measurement systems were obtained and processed as part of an exploratory CART analysis. Data for the precursor species sulfur dioxide (SO_2) and oxides of nitrogen (NO_x) were also obtained and processed for two of the areas of interest.

2.1.1. Data Sources and Initial Processing Steps

All air quality data were obtained from the AIRS (Atmospheric Information Retrieval System; http://www.epa.gov/air/data/index.html) database. This database is updated regularly by EPA and the latest version of the database at the time of data retrieval was used.

In preparing the $PM_{2.5}$ data, we first identified all monitoring sites within and potentially upwind of each area of interest and determined whether the data for each individual site are FRM, continuous, or speciated. We also determined the data collection interval. For upwind sites, we required the availability of both daily FRM and continuous data—the former for use in the CART analysis as an indicator of the prior day's upwind $PM_{2.5}$ concentration and the latter for use in forecasting. We then extracted and reformatted the FRM data for each available site. For most sites, the FRM data are available on a daily basis. For two areas, Bristol and Roanoke, Virginia, the FRM data are available every three days.

During the course of the $PM_{2.5}$ forecasting project, several exploratory CART analyses were performed that used additional air quality data. SO_2 and NO_x data were obtained from AIRS and processed for sites in the Baltimore and Charlotte areas. Continuous $PM_{2.5}$ data were also obtain and processed for all of the areas of interest and associated upwind areas.

To ensure the reliability of the underlying data from the AIRS database as well as the extraction and reformatting steps, we conducted the following quality assurance checks for all data:

- State and county codes for each site were verified.
- Units for all data elements were confirmed.
- Randomly selected values in the re-formatted files were cross-checked against the original data files for accuracy.
- $PM_{2.5}$ (or other species) values for each site were extracted and sorted according to magnitude, to check the range of values for reasonableness (e.g., that all concentration values are positive) and the completeness of the dataset (i.e., that missing values are accounted for and properly indicated).

2.1.2. Summary of Data Sites and Parameters

Table 2–1 lists the air quality data sites for each area of interest, as used for the MARAMA $PM_{2.5}$ forecasting tool development project. Both local and potential upwind sites are listed; only local sites were used in determining the area-wide maximum $PM_{2.5}$ concentration for input to CART.

Table 2-1. Summary of Air Quality Monitoring Sites and Data Used in the MARAMA $PM_{2.5}$ Forecasting Tool Development Project

Area/Site Name	AIRS ID	Pollutant	Measurement Type	Sampling Frequency	Date Commenced *(for period of study)*	Use in CART Analysis
Charlotte						
Kannapolis	370250004	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Gastonia	370710016	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Charlotte #10 Fire Station	371190010	$PM_{2.5}$	FRM	Daily	1/99	Local
Charlotte Plaza	371190034	$PM_{2.5}$	FRM	Daily	1/99–7/99	Local
Charlotte #16 Fire Station	371190040	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Charlotte Garinger	371190041	$PM_{2.5}/SO_2/NO_x$	FRM/TEOM/ Analyzers	Daily/Hourly	7/99	Local/ Recirculation/ Exploratory
Emerywood Dr.	371190042	$PM_{2.5}$	FRM	1 in 3 days	9/00	Local
HWY 321—Back Field	450910006	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Winston-Salem	370670022	$PM_{2.5}/SO_2$	FRM	Daily/Hourly	1/99	Upwind/ Exploratory
Greenville	450450009	$PM_{2.5}$	FRM	Daily		Upwind
Spartanburg	450450010	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Greenville	450450008	SO_2	Analyzer	Hourly		Exploratory
Bristol						
Sullivan Co, TN	471631007	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Highlands View Elementary School	515200006	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Knoxville—Davanna St.	470931013	$PM_{2.5}$	TEOM	Daily	1/99	Upwind
Knoxville—Vermont Ave.	470931017	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Knoxville—Mildred Dr.	470931020	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Roanoke						
Raleigh Court Library	517700014	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Market Street Fire Station	517750010	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local

Area/Site Name	AIRS ID	Pollutant	Measurement Type	Sampling Frequency	Date Commenced *(for period of study)*	Use in CART Analysis
Winston-Salem	370670022	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Richmond						
Shirley Plantation	510360002	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Bensley Armory	510410003	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Mathematics & Science Center	510870014	$PM_{2.5}$	FRM/TEOM	Daily	1/99	Local/ Recirculation
DEQ Regional Office	510870015	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
DEQ Air Monitoring Office	517600020	$PM_{2.5}$	FRM	Daily	1/99	Local
McMillan/DC	110010043	$PM_{2.5}$	FRM/TEOM	Daily	1/99	Upwind
Winston-Salem	370670022	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Washington, D.C.						
River Terrace School	110010041	$PM_{2.5}$	FRM	Daily	2/99	Local
Ohio Drive	110010042	$PM_{2.5}$	FRM	1 in 3 days	2/99	Local
McMillan Reservoir	110010043	$PM_{2.5}$	FRM	Daily	1/99	Local/ Recirculation
Rockville	240313001	$PM_{2.5}$	FRM	1 in 3 days	7/99	Local
Goddard Space Center	240330002	$PM_{2.5}$	FRM	1 in 3 days	7/02	Local
Suitland	240338001	$PM_{2.5}$	FRM	1 in 3 days	8/99	Local
Aurora Hills Vis. Ctr.	510130020	$PM_{2.5}$	FRM	1 in 3 days		Local
Lee District Park	510290030	$PM_{2.5}$	FRM	Daily	1/99	Local
Steven Corners	510591004	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Lewinsville	510595001	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Broad Run High School	511071005	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Old Town (Baltimore)	245100040	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Gettysburg	420010001	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Math & Science Center (Richmond)	510870014	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Baltimore						
Davidsonville	240030014	$PM_{2.5}$	FRM	1 in 3 days	8/99	Local
Ft. Meade	240030019	$PM_{2.5}$	FRM	1 in 3 days	2/99	Local
Glen Burnie	240031003	$PM_{2.5}$	FRM	1 in 3 days	11/99	Local
Riviera Beach	240032002	$PM_{2.5}/SO_2$	FRM/ Analyzer	1 in 3 days	2/99	Local/ Exploratory

Area/Site Name	AIRS ID	Pollutant	Measurement Type	Sampling Frequency	Date Commenced *(for period of study)*	Use in CART Analysis
Padonia	240051007	$PM_{2.5}$	FRM	1 in 3 days	1/00	Local
Essex	240053001	$PM_{2.5}$	FRM	Daily	8/99	Local
Edgewood	240251001	$PM_{2.5}$	FRM	1 in 3 days	8/99	Local
NEPS	245100006	$PM_{2.5}$	FRM	1 in 3 days	8/99	Local
NWPS	245100007	$PM_{2.5}$	FRM	1 in 3 days	8/99	Local
SE Police Station	245100008	$PM_{2.5}$	FRM	1 in 3 days	6/01	Local
FMC	245100035	$PM_{2.5}$	FRM	Daily	8/99	Local
Old Town	245100040	$PM_{2.5}$/ NO_x	FRM/TEOM	Daily	1/99	Local/ Recirculation/ Exploratory
Westport	245100049	$PM_{2.5}$	FRM	1 in 3 days		Local
Fire Stn. #50	245100052	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
McMillan Reservoir (Washington)	110010043	$PM_{2.5}$	FRM	Daily	1/99	Recirculation
Gettysburg	420010001	$PM_{2.5}$	FRM	Daily	1/99	Recirculation
Math & Sci. Center (Richmond)	517600020	$PM_{2.5}$	FRM	Daily	1/99	Upwind
River Terrace School	110010041	SO_2	FRM	Hourly	2/99	Exploratory
Sci. Museum	517600024	SO_2	FRM	Hourly	1/99	Exploratory
Philadelphia						
AMS Lab	421010004	$PM_{2.5}$	FRM	Daily	2/99	Local
Belmont Water Treatment	421010020	$PM_{2.5}$	FRM	1 in 3 days	3/99	Local
Northeast Airport	421010024	$PM_{2.5}$	FRM	1 in 3 days	2/99	Local
Community Health Services	421010047	$PM_{2.5}$	FRM	1 in 3 days	2/99	Local
Elmwood	421010136	$PM_{2.5}$	FRM	Daily	2/99	Local
Roxy Water Pump	421010014	$PM_{2.5}$	FRM	1 in 3 days		Exploratory
Camden Lab	340070003	$PM_{2.5}$	FRM	1 in 3 days		Exploratory
Pennsauken	340071007	$PM_{2.5}$	FRM	1 in 3 days		Exploratory
Gibbstown	340155001	$PM_{2.5}$	FRM	1 in 3 days		Exploratory
Bristol	420170012	$PM_{2.5}$	FRM	1 in 3 days		Exploratory
Chester	420450002	$PM_{2.5}$	FRM	1 in 3 days		Exploratory
Norristown	420910013	$PM_{2.5}$	FRM	1 in 3 days		Exploratory

Area/Site Name	AIRS ID	Pollutant	Measurement Type	Sampling Frequency	Date Commenced *(for period of study)*	Use in CART Analysis
McMillan Reservoir (Washington)	110010043	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Old Town (Baltimore)	245100040	$PM_{2.5}$ /NO_x	FRM	Daily	1/99	Upwind
Gettysburg	420010001	$PM_{2.5}$	FRM	Daily	1/99	Upwind
New Castle—MLK	100032004	$PM_{2.5}$	FRM	Daily	2/99	Recirculation
Camden	340070003	$PM_{2.5}$	FRM/TEOM	Daily	1/99	Recirculation
Wilmington						
Bellefonte	100031003	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Lums Pond	100031007	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Newark UD	100031011	$PM_{2.5}$	FRM	1 in 3 days	3/99–11/99	Local
Newark	100031012	$PM_{2.5}$	FRM	1 in 3 days	12/99	Local
New Castle—MLK	100032004	$PM_{2.5}$	FRM	Daily	2/99	Local/ Recirculation
Fairhill	240150003	$PM_{2.5}$	FRM	1 in 3 days	11/99	Local
McMillan Reservoir (Washington)	110010043	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Old Town (Baltimore)	245100040	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Gettysburg	420010001	$PM_{2.5}$	FRM	Daily	1/99	Upwind
Newark						
Fort Lee	340030003	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Newark	340130011	$PM_{2.5}$	FRM/TEOM	1 in 3 days	1/99	Local
Willis Center	340130015	$PM_{2.5}$	FRM	1 in 3 days	4/99	Local
Lexington	340130016	$PM_{2.5}$	FRM	1 in 3 days	7/01	Local
Ryders Lane	340230006	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Elizabeth Lab	340390004	$PM_{2.5}$	FRM	Daily	1/99	Local/ Recirculation
Elizabeth—Mitchell	340390006	$PM_{2.5}$	FRM	1 in 3 days	1/99	Local
Rahway	340392003	$PM_{2.5}$	FRM	1 in 3 days	12/99	Local
MLK (New Castle)	100032004	$PM_{2.5}$	FRM/TEOM	Daily	2/99	Upwind
Camden	340070003	$PM_{2.5}$	FRM/TEOM	Daily	1/99	Upwind
Bethelehem-Freemansburg	420950025	$PM_{2.5}$	FRM	Daily	1/99	Upwind

Only the entries labeled local and upwind were used in the final CART analyses. The data for the local sites were used to calculate the daily maximum $PM_{2.5}$ concentration for the areas of interest. The data for the upwind sites were used to provide information about possible transport or recirculation of PM. For each area of interest with more than one local $PM_{2.5}$ monitoring site, the maximum over all local sites was determined and used to represent the daily $PM_{2.5}$ concentration for that area. Similarly, for upwind areas with more than one $PM_{2.5}$ monitoring site, the maximum over all sites was used. In the exploratory analyses, data for the individual sites were used independently.

The local $PM_{2.5}$ concentration for each area provided the classification parameter for the CART analysis. Specifically, the classification variable for each area was assigned a value of 1, 2 or 3 based on the value of the local daily maximum concentration. Each classification category represents a different range of PM2.5 concentration. The three categories were defined based on the EPA established guidelines for PM2.5 forecasting as follows: less than 15.5 (Category 1), 15.5 to less than 40.5 (Category 2), and greater or equal to 40.5 μgm^{-3} (Category 3). Since only a few data points were in the highest EPA category of greater than or equal to 65 μgm^{-3}, this category was not used in the analysis. The three categories used in this analysis are also referred to by the colors: green, yellow, and orange and by the descriptors “good”, “moderate”, and “unhealthy for sensitive groups (USG).”

The specific air quality parameters used in the final CART analysis for each area are listed and described in Table 2-2. In this table and throughout the discussion of the CART analysis, the “analysis” day is the day that is classified by CART and the two-days-prior day is the day two days prior to the analysis day. Note that later in the report, the “analysis” day is the “forecast” day. In both cases, it is the day for which the classification analysis or the forecast is being made.

Table 2-2. Summary of $PM_{2.5}$ Parameters Used in the Final CART Analysis to Support the MARAMA $PM_{2.5}$ Forecasting Tool Development

Forecast Area	Parameter Name	Description	Units
Charlotte			
	bpm_c	*The classification parameter.* It has a value of 1, 2, or 3 such that each value corresponds to a 24-hour average PM2.5 concentration for the analysis day that is <15.5, 15.5 to < 40.5, or ≥ 40.5 μgm^{-3}.	none
	y2dpm_gs	The 24-hour average PM2.5 concentration for two days prior at Greenville-Spartanburg.	μgm^{-3}
	y2dpm_me	The 24-hour average PM2.5 concentration for two days prior at Mecklenberg.	μgm^{-3}
	y2dpm_ws	The 24-hour average PM2.5 concentration for two days prior at Winston-Salem.	μgm^{-3}
Bristol			
	bpm_br	*The classification parameter.* It has a value of 1, 2, or 3 such that each value corresponds to a 24-hour average PM2.5 concentration for the analysis day that is <15.5, 15.5 to < 40.5, or ≥ 40.5 μgm^{-3}.	none
	y2dpm_kn	The 24-hour average PM2.5 concentration for two days prior at Knoxville.	μgm^{-3}
Roanoke			

Forecast Area	Parameter Name	Description	Units
	bpm_ro	*The classification parameter.* It has a value of 1, 2, or 3 such that each value corresponds to a 24-hour average PM2.5 concentration for the analysis day that is <15.5, 15.5 to < 40.5, or ≥ 40.5 μgm^{-3}.	none
	y2dpm_ws	The 24-hour average PM2.5 concentration for two days prior at Winston-Salem.	μgm^{-3}
Richmond			
	bpm_r	*The classification parameter.* It has a value of 1, 2, or 3 such that each value corresponds to a 24-hour average PM2.5 concentration for the analysis day that is <15.5, 15.5 to < 40.5, or ≥ 40.5 μgm^{-3}.	none
	y2dpm_mc	The 24-hour average PM2.5 concentration for two days prior at Washington D.C. (McMillian).	μgm^{-3}
	y2dpm_rh	The 24-hour average PM2.5 concentration for two days prior at Richmond.	μgm^{-3}
	y2dpm_ws	The 24-hour average PM2.5 concentration for two days prior at Winston-Salem.	μgm^{-3}
Washington D.C.			
	bpm_dc	*The classification parameter.* It has a value of 1, 2, or 3 such that each value corresponds to a 24-hour average PM2.5 concentration for the analysis day that is <15.5, 15.5 to < 40.5, or ≥ 40.5 μgm^{-3}.	none
	mxmcotgy	The maximum 24-hour average PM2.5 concentration for two days prior at McMillian (Washington D.C), Old Town (Baltimore), and Gettysburg.	μgm^{-3}
	y2dpm_rh	The 24-hour average PM2.5 concentration for two days prior at Richmond.	μgm^{-3}
Baltimore			
	bpm_b	*The classification parameter.* It has a value of 1, 2, or 3 such that each value corresponds to a 24-hour average PM2.5 concentration for the analysis day that is <15.5, 15.5 to < 40.5, or ≥ 40.5 μgm^{-3}.	none
	mxmcotgy	The maximum 24-hour average PM2.5 concentration for two days prior at McMillian (Washington D.C), Old Town (Baltimore), and Gettysburg.	μgm^{-3}
	y2dpm_rh	The 24-hour average PM2.5 concentration for two days prior at Richmond.	μgm^{-3}
Philadelphia			
	bpm_p	*The classification parameter.* It has a value of 1, 2, or 3 such that each value corresponds to a 24-hour average PM2.5 concentration for the analysis day that is <15.5, 15.5 to < 40.5, or ≥ 40.5 μgm^{-3}.	none
	mxcanw	The maximum 24-hour average PM2.5 concentration for two days prior at Camden and New Castle.	μgm^{-3}
	mxmcotgy	The maximum 24-hour average PM2.5 concentration for two days prior at McMillian (Washington D.C), Old Town (Baltimore), and Gettysburg.	μgm^{-3}
Wilmington			
	bpm_w	*The classification parameter.* It has a value of 1, 2, or 3 such that each value corresponds to a 24-hour average PM2.5 concentration for the analysis day that is <15.5, 15.5 to < 40.5, or ≥ 40.5 μgm^{-3}.	none
	mxmcotgy	The maximum 24-hour average PM2.5 concentration for two days prior at McMillian (Washington D.C), Old Town (Baltimore), and Gettysburg.	μgm^{-3}

Forecast Area	Parameter Name	Description	Units
	y2dpm_nw	The 24-hour average PM2.5 concentration for two days prior at New Castle.	μgm^{-3}
Newark			
	bpm_n	*The classification parameter.* It has a value of 1, 2, or 3 such that each value corresponds to a 24-hour average PM2.5 concentration for the analysis day that is <15.5, 15.5 to < 40.5, or ≥ 40.5 μgm^{-3}.	none
	mxcanw	The maximum 24-hour average PM2.5 concentration for two days prior at Camden and New Castle.	μgm^{-3}
	y2dpm_ez	The 24-hour average PM2.5 concentration for two days prior at Elizabeth.	μgm^{-3}

In the data files that accompany this report, the site-specific portions of the parameter names are defined as follows:

gs = Greenville-Spartanburg, SC

me = Charlotte, NC (Mecklenburg Co.)

ws = Winston-Salem, NC

kn = Knoxville, TN

rh = Richmond, VA

mc = Washington, D.C. (McMillan Reservoir)

ot = Baltimore, MD (Old Town)

nw = New Castle, DE

ca = Camden, NJ

ez = Elizabeth, NJ

gy = Gettysburg, PA

2.1.3. Problems and Limitations

A key limitation of the study is related to the availability of historical $PM_{2.5}$ data for use in the CART analysis. As indicated in Table 2-1, $PM_{2.5}$ monitoring began during 1999 or 2000 for most sites/areas and data completeness ranged from approximately 65 to 100 percent for the dependent variable, based on the full period of 1999–2002. For the Bristol and Roanoke sites in Virginia, data are available only every three days. Use of data for a three- to four-year period of record with few high $PM_{2.5}$ values may limit the ability of CART to identify the key high $PM_{2.5}$ regimes or distinguish the complete set of conditions that lead to the various $PM_{2.5}$ levels—simply because the high PM days and/or the full range of meteorological conditions are not represented by a sufficient number of days in the historical database.

An important consideration in the use of the historical data to develop a real-time forecasting tool is, of course, the availability of real-time data to support the forecasting. $PM_{2.5}$ data collected using the FRM measurement systems were used for the CART analysis—as they are expected to provide the most consistent and accurate concentration values. It follows that these data are best suited for establishing meaningful relationships between meteorological

parameters and PM concentration. However, because they are collected using filters, data are typically not available until several weeks after the sampling date. Instead, forecasters must rely on continuous measurements of $PM_{2.5}$ (which are available on a real time basis) to provide information about prior day PM levels at local and upwind sites and to support the forecasting. There are several different types of instruments used to collect continuous data, and these do not always agree with the FRM measurements. The level of disagreement varies from site to site, and typically from season to season (with temperature and humidity), as discussed in some detail by Gillespie et al. (2004). The issue for the CART-based forecasting project is that the real-time data from continuous measurement systems may be different enough from the FRM data under some circumstances to cause an erroneous forecast. For most areas, prior day PM2.5 concentrations were important to the CART analysis and thus to the forecasts - increasing the possibility that differences in the data types could contribute to forecast errors. Adjusting the continuous data to an FRM equivalent value is an option for the forecasters to use to overcome this limitation.

2.2. Meteorological Data

The meteorological data used for this study consist of measurements of various surface and upper-air meteorological parameters for sites located within and nearby each area of interest. To represent the local- and regional-scale meteorological conditions for each area, we selected one local surface meteorological monitoring site and one or more nearby upper-air monitoring site(s). Upper-air data collected using profiler measurement systems were also obtained and processed for several areas as part of an exploratory analysis.

2.2.1. Data Sources and Initial Processing Steps

The historical surface and upper-air data meteorological data were obtained from the National Climatic Data Center (NCDC), either via the Internet or from published CD databases. Profiler data were obtained from the National Oceanic and Atmospheric Administration (NOAA).

To ensure the reliability of the meteorological data as well as the extraction and reformatting steps, we conducted the following quality assurance checks for all data:

- All source codes used to collect and reprocess data from the original format to that used by CART were specifically reviewed before application to confirm the suitability of the data processing software for the data type/format.
- The units for all data elements and for all sites were confirmed.
- The range of time over which the data are available and the time stamp for each data element were reviewed.
- For data elements that are used directly by CART, several (at least ten) random dates and times were selected and the values of the meteorological data elements were spot-checked against the original data files.
- For data elements that are computed from the original values, several (at least 10) random dates and times were selected and the values of the derived quantities were checked against independent calculations using the original data.
- The values of the meteorological parameters for each site were sorted according to magnitude, to check the range of values for reasonableness (e.g., that all values are within

expected ranges for each parameter) and the completeness of the dataset (i.e., that missing values are accounted for and properly indicated).

2.2.2. Summary Tables

Table 2-3 lists the surface meteorological data sites for each area of interest, as used for the MARAMA $PM_{2.5}$ forecasting tool development project.

Table 2-3. Summary of Surface Meteorological Monitoring Sites and Data Used in the MARAMA $PM_{2.5}$ Forecasting Tool Development Project

Area/Site Name	WBAN Number	Parameters*	Sampling Frequency	Availability During 1999–2002
Charlotte				
CLT—Charlotte Douglas Intl. Airport	13881	T, RH, WS, WD, Precip	Hourly	1/99–12/02
Bristol				
TRI—Bristol Tri Cities Airport	13877	T, RH, WS, WD, Precip	Hourly	1/99–12/02
Roanoke				
ROA—Roanoke Regional Airport	13741	T, RH, WS, WD, Precip	Hourly	1/99–12/02
Richmond				
RIC—Richmond International Airport	13740	T, RH, WS, WD, Precip	Hourly	1/99–12/02
Washington, D.C.				
DCA—Washington Regan National Airport	13743	T, RH, WS, WD, Precip	Hourly	1/99–12/02
IAD—Washington D.C. Dulles Intl. Airport	93738	WS, WD	Hourly	1/99–12/02
Baltimore				
BWI - Baltimore Washington Intl. Airport	93721	T, RH, WS, WD, Precip	Hourly	1/99–12/02
Philadelphia				
PHL - Philadelphia Intl. Airport	13739	T, RH, WS, WD, Precip	Hourly	1/99–12/02
Wilmington				
Wilmington New Castle County Airport	13781	T, RH, WS, WD, Precip	Hourly	1/99–12/02
Newark				
EWR—Newark Intl. Airport	14734	T, RH, WS, WD, Precip	Hourly	1/99–12/02

**T= temperature, RH = relative humidity, WS = wind speed, WD = wind direction…*

The NWS surface meteorological datasets were largely complete. Missing data were appropriately flagged in the CART datasets.

The upper-air meteorological data sites for each area of interest, as used for the MARAMA $PM_{2.5}$ forecasting tool development project are listed and summarized in Table 2-4. The upper-air monitoring sites were matched to the areas of interest based on proximity and in an attempt to best represent the regional airflow patterns within the surrounding area. Location relative to geographic features, including the coastline, was also considered.

Table 2-4. Summary of Upper-Air Meteorological Monitoring Sites and Data Used in the MARAMA $PM_{2.5}$ Forecasting Tool Development Project

Area/Site Name	WBAN Number	Parameters*	Sampling Frequency	Availability During 1999–2002
Charlotte				
GSO—Greensboro	13723	T, RH-Cloud, WS, WD, ϕ	Twice per day (0 and 12Z)	1/99–12/02
Bristol				
RNK—Roanoke/Blackburg	53829	T, RH-Cloud, WS, WD, ϕ	Twice per day (0 and 12Z)	1/99–12/02
Roanoke				
RNK—Roanoke/Blackburg	53829	T, RH-Cloud, WS, WD, ϕ	Twice per day (0 and 12Z)	1/99–12/02
Richmond				
IAD—Sterling/Washington D.C./Dulles	93734	T, RH-Cloud, WS, WD, ϕ	Twice per day (0 and 12Z)	1/99–12/02
Washington, D.C.				
IAD—Sterling/Washington D.C./Dulles	93734	T, RH-Cloud, WS, WD, ϕ	Twice per day (0 and 12Z)	1/99–12/02
Baltimore				
IAD—Sterling/Washington D.C./Dulles	93734	T, RH-Cloud, WS, WD, ϕ	Twice per day (0 and 12Z)	1/99–12/02
Philadelphia				
IAD—Sterling/Washington D.C./Dulles	93734	T, RH-Cloud, WS, WD, ϕ	Twice per day (0 and 12Z)	1/99–12/02
Wilmington				
IAD—Sterling/Washington D.C./Dulles	93734	T, RH-Cloud, WS, WD, ϕ	Twice per day (0 and 12Z)	1/99–12/02
Newark				
OKX—Broohkaven	94793	T, RH-Cloud, WS, WD, ϕ	Twice per day (0 and 12Z)	1/99–12/02

**T= temperature, RH-Cloud = cloud index based on relative humidity, WS = wind speed, WD = wind direction, ϕ = geopotential height*

The specific surface meteorological parameters used in the final CART analysis for each area are listed and described in Table 2-5. In this table and throughout the discussion of the CART analysis, the “analysis” day is the day that is classified by CART and the prior day is the day prior to the analysis day. Note that later in the report, the “analysis” day is the “forecast” day. In both cases, it is the day for which the classification analysis or the forecast is being made.

Table 2-5. SUMMARY of Surface Meteorological Parameters Used in the Final CART Analysis to Support the MARAMA PM2.5 Forecasting Tool Development

Variable names are generic and vary slightly for each monitoring site.

Parameter Name	Description	Units
tmax_xx	Daily maximum surface temperature for the analysis day.	°C
tmin_xx	Daily minimum surface temperature for the analysis day.	°C
rh24_xx	Average relative humidity for the analysis day based on temperatures and dew point temperatures at 6Z, 9Z, 12Z, 15Z, 18Z, 21Z, 0Z, 3Z	%
pflg4_xx	Number of 6-hourly periods with rainfall greater that 0.1 inches for the analysis day.	unitless (value of 0–4)
wb24_xx	Average (vector) wind direction bin for the analysis day based on values at 6Z, 9Z, 12Z, 15Z, 18Z, 21Z, 0Z, 3Z . Binned such that 1= N, 2 = E, 3 = S, 4 = W, 5 = Calm). Not used for Philadelphia, Wilmington, or Newark.	unitless (value of 1–5)
wb24_xx2	Average (vector) wind direction bin for the analysis day based on values at 6Z, 9Z, 12Z, 15Z, 18Z, 21Z, 0Z, 3Z. Binned such that 1= NE, 2 = SE, 3 = SW, 4 = NW, 5 = Calm). Used for Philadelphia, Wilmington, and Newark.	unitless (value of 1–5)
ws24_xx	Average (vector) wind speed for the analysis day based on values at 6Z, 9Z, 12Z, 15Z, 18Z, 21Z, 0Z, 3Z.	ms^{-1}

In Table 2-5 and data files that accompany this report, the site-specific portions (*xx*) of the parameter names are defined as follows:

c = Charlotte, NC

br = Bristol, VA

ro = Roanoke, VA

r = Richmond, VA

dc = Washington, D.C. (Reagan/National Airport)

d = Washington, D.C. (Dulles Airport)

b = Baltimore, MD (Old Town)

p = Philadelphia, PA

w = Wilmington, DE

ne = Newark, NJ

The upper-air meteorological parameters are listed and described in Table 2-6.

Table 2-6. Summary of Upper-Air Meteorological Parameters Used in the Final CART Analysis to Support the MARAMA PM2.5 Forecasting Tool Development

Variable names are generic and vary slightly for each monitoring site.

Parameter Name	Description	Units
t85amxx	850 mb temperature corresponding to the morning sounding (12Z) on the analysis day.	°C
t85pmx	850 mb temperature corresponding to the evening sounding (0Z) on the analysis day.	°C
delt950x	Difference in temperature between the 950 mb temperature corresponding to the morning sounding (12Z) on the analysis day and that at the surface of the same sounding. Not used for Bristol or Roanoke.	°C
delt900x	Difference in temperature between the 900 mb temperature corresponding to the morning sounding (12Z) on the analysis day and that at the surface of the same sounding.	°C
delt850x	Difference in temperature between the 850 mb temperature corresponding to the morning sounding (12Z) on the analysis day and that at the surface of the same sounding.	°C
htthty7x	Height difference computed as the difference of the average 700 mb geopotential height on the current day (corresponding to the morning and evening soundings) and the average 700 mb geopotential height of the day prior to the analysis day (corresponding to the morning and evening soundings of that day).	m
cloudx	Cloud index defined as the maximum of the cloud indexes determined from the relative humidity of the morning and evening soundings. CLOUDAMx = equal to1, 2, or 3 as follows, using the relative humidity (RH) corresponding to the morning (AM) sounding at 850 mb and 700 mb = 1 if RH 850 AM < 80 and RH 700 AM <65 = 2 if RH 850 AM >=80 and RH 700 AM <65 = 2 if RH 850 AM < 80 and RH 700 PM >=65 =3 if RH 850 AM >=80 and RH 700 AM >=65 CLOUDPMx = equal to1, 2, or 3 as follows, using the relative humidity (RH) corresponding to the morning (AM) sounding at 850 mb and 700 mb = 1 if RH 850 PM < 80 and RH 700 PM <65 = 2 if RH 850 PM >=80 and RH 700 PM <65 = 2 if RH 850 PM < 80 and RH 700 PM >=65 =3 if RH 850 PM >=80 and RH 700 PM >=65 Cloudx is then the maximum of cloudamx and cloudpmx	none
ywb85pmx	850 mb wind direction corresponding to the evening sounding (0Z) of the day prior to the analysis day. Binned such that 1= N, 2 = E, 3 = S, 4 = W, 5 = Calm).	unitless (value of 1–5)
ywb85pmx2	850 mb wind direction corresponding to the evening sounding (0Z) of the day prior to the analysis day. Binned such that 1= NE, 2 = SE, 3 = SW, 4 =NW, 5 = Calm). Used for Philadelphia, Wilmington, and Newark.	unitless (value of 1–5)

Parameter Name	Description	Units
ywb70pmx	700 mb wind direction bin corresponding to the evening sounding (0Z) of the day prior to the analysis day (binned such that 1= N, 2 = E, 3 = S, 4 = W, 5 = Calm). Not used for Philadelphia, Wilmington, and Newark.	unitless (value of 1–5)
ywb70pmx2	700 mb wind direction bin corresponding to the evening sounding (0Z) of the day prior to the analysis day (binned such that 1= NE, 2 = SE, 3 = SW, 4 = NW, 5 = Calm). Used for Philadelphia, Wilmington, and Newark.	unitless (value of 1–5)
wb85amx	850 mb wind direction bin corresponding to the morning sounding (12Z) of the analysis day. Binned such that 1= N, 2 = E, 3 = S, 4 = W, 5 = Calm. Not used for Philadelphia, Wilmington, and Newark.	unitless (value of 1–5)
wb85amx2	850 mb wind direction corresponding to the morning sounding (12Z) of the analysis day. Binned such that 1= NE, 2 = SE, 3 = SW, 4 = NW, 5 = Calm. Used for Philadelphia, Wilmington, and Newark.	unitless (value of 1–5)
wb85pmx	850 mb wind direction bin corresponding to the evening sounding (0Z) of the analysis day. Binned such that 1= N, 2 = E, 3 = S, 4 = W, 5 = Calm. Not used for Philadelphia, Wilmington, and Newark.	unitless (value of 1–5)
wb85pmx2	850 mb wind direction bin corresponding to the evening sounding (0Z) of the analysis day. Binned such that 1= NE, 2 = SE, 3 = SW, 4 = NW, 5 = Calm. Used for Philadelphia, Wilmington, and Newark.	unitless (value of 1–5)
vawb85x	850 mb vector average wind direction determined from morning (12 Z) and evening (0 Z) soundings at 850 mb (binned such that 1= NE, 2 = SE, 3 = S, W 4 = NW, 5 = Calm). Used for Charlotte (Roanoke).	unitless (value of 1–5)
yws85pmx	850 mb wind speed corresponding to the evening sounding (0Z) of the day prior to the analysis day	ms-1
yws70pmx	700 mb wind speed corresponding to the evening sounding (0Z) of the day prior to the analysis day	ms-1
ws85amx	850 mb wind speed corresponding to the morning sounding (12Z) of the analysis day	ms-1
ws85pmx	850 mb wind speed corresponding to the evening sounding (0Z) of the analysis day	ms-1
vaws85x	850 mb vector average wind speed determined from morning (12 Z) and evening (0 Z) soundings on the analysis day. Used for Charlotte (Roanoke).	ms-1

In Table 2-6 and data files that accompany this report, the site-specific (*xx*) portions of the parameter names are defined as follows:

G = Greensboro, NC

R = Roanoke, VA

D = Washington, D.C. (Dulles Airport)

B = Brookhaven, NY

In addition to the surface and upper-air meteorological data and the air quality data, an additional variable, seas3 was used. This variable was set equal to “1” if the analysis day was in the month of January, February, March, November, or December. The variable was set equal to “2” if the analysis day was in the month of April, May, September, or October. And lastly was set equal to “3” if the analysis day was in the month of June, July, or August.

2.2.3. Problems and Limitations

For this analysis we used primarily routine NWS data and data quality and completeness was generally very good. We encountered one issue with the surface data. For Washington D.C., the surface wind data for the Dulles Airport monitor were substituted during the course of the analysis for the surface wind data for Reagan/National Airport. Although the Reagan/National Airport is located closer to the urban area, the location of the wind monitor relative to the Potomac River is expected to cause the winds from this monitor to be unrepresentative of the area. Thus the surface winds for the Dulles Airport monitor, located in an open area to the west of the city, were used instead.

One issue regarding the use of the upper-air data is that with the exception of Roanoke and Washington, D.C., there are no upper-air monitoring sites within the areas of interest. Thus we were required to use data for the nearest upper-air monitoring sites to describe the upper-air conditions. The assignments, as given in Table 2-4, were based on proximity. Location relative to geographic features, including the coastline, was also considered. In general, good matches were achieved with either nearby or similarly located sites. Nevertheless, the lack of local upper-air data is a limitation for the analysis.

One possible solution to the lack of local upper-air data is the use of profiler data, where available. As part of this study, we investigated the use of profiler data for Baltimore using the Ft. Meade profiler data. Because of a lack of moisture measurements and temperature data coupled in time/space, we used only the wind data available from Ft. Meade. Moisture, temperature, and geopotential height data for the CART simulations were based on measurements from the Dulles soundings for the CART runs. We found that the results using the Ft. Meade wind data were similar to those resulting from the use of Dulles sounding data. Since nothing appeared to be gained from the use of these data, the use of the more standard, readily available data from Dulles was chosen.

Key issues with the use of the profiler data were that moisture data were not available and temperature data were either not available, or were not coupled in time. Also, for the most part, data at the sites of interest were not available for the entire analysis period.

2.3. Electronic Datasets

The CART input datasets for each area are provided as an electronic attachment to this report (Attachment A). The air quality data were processed using Microsoft ACCESS and EXCEL on personal computers (PCs). The meteorological data were initially processed using UNIX Fortran programs on main-frame computers and the data were then passed to PCs where they were converted to EXCEL format. The air quality and meteorological data from the various sources were then merged into EXCEL spreadsheets for each area of interest. These data files were then converted into systat (*.sys) format using DBMS/Copy for Windows. It was at this point that additional “computed” parameters were added (i.e. for cloud, season, maximum PM), final missing data were set/flagged consistently, the databases were “stripped” of days not meeting criteria for a given area, and final QA/QC was performed. Days with missing dependent variables were not specifically stripped out and as a result the final CART-ready databases do contain days with missing dependent variables. CART itself handles these days appropriately. CART was run using these final *.sys formatted files. Data files provided to the MARAMA participants are EXCEL files that have been created from these final *.sys files used in CART. All blank (missing) cells have been replaced with “-999.”

This page intentionally left blank.

3. CART Analysis Methods and Results

In this section we discuss the application of CART for each of the areas of interest. We begin with a brief overview of the CART program.

3.1. Overview of CART

The CART analysis software (Brieman et al., 1984; Steinberg et al., 1997) is a statistical analysis tool that partitions a dataset into discrete subgroups based on the value of a user-defined classification variable (e.g., 24-hour average $PM_{2.5}$ concentration). The remaining variables in the database are selected as to whether or not they provide a segregation of the data for different values of the classification variable. The analysis procedure assumes that there is a causal relationship between the independent variables and the dependent variable. Consequently, it is necessary to construct a database of independent variables such that this relationship can be identified.

The CART technique is designed to segregate objects or, in the case of air quality analysis, days with different values of a classification variable into different bins or terminal nodes. The CART technique accomplishes this task through the development of a binary decision tree, comprised of a progression of binary splits on the values of the independent variables. At each split, or node, the data are divided according to their value for one of the independent variables, in a way that improves their segregation by the dependent variable. The end of a branch—called a terminal node, or bin—corresponds to a subset of the data with predominantly one value for the classification variable, characterized by independent variable ranges defined along the path to that bin. Thus the tree identifies parameter conditions frequently associated with values of the dependent variable. The user specifies the desired complexity of the tree, that is, the degree of branching and resulting number of terminal bins.

The parameter and parameter values associated with the CART classification tree provide information on the relative importance of the various air quality and meteorological parameters to the air quality conditions as represented by the dependent variable. Thus the CART technique not only segregates the days, but does so in a manner that provides physical insight into the classified days. This physical insight allows the analyst to examine whether the data partitioning is meaningful.

By segregating the data values into the classification bins, CART also provides information regarding the frequency of occurrence of the conditions associated with each classification category. In this manner, the likely recurrence rate for a particular type of day and the associated prevailing conditions are obtained.

3.2. CART Application Procedures

The primary goal of this project was to use the results of the CART application to develop a forecasting algorithm for each area of interest. CART was applied for a multi-year period (nominally 1999–2002) and all days with available data within this period were classified and grouped into bins in accordance with the values of observed and calculated meteorological and air quality parameters that comprise the input dataset.

As discussed in detail in Section 2 of this report, we used 24-hour average $PM_{2.5}$ concentration as the classification variable for this application. The classification variable for each area was assigned a value of 1, 2 or 3 based on the value of the local daily maximum concentration. The

categories were defined based on the EPA established guidelines for PM2.5 forecasting as follows: less than 15.5 (Category 1), 15.5 to less than 40.5 (Category 2), and greater or equal to 40.5 μgm^{-3} (Category 3). Since only a few data points were in the highest EPA category of greater than or equal to 65 μgm^{-3}, this category was not used in the analysis. In applying CART we also included a variety of meteorological and air quality parameters as input data, as discussed in detail in Section 2. CART was applied separately for each area of interest.

CART requires the specification of "costs" associated with the misclassification of days into bins corresponding to a different category than indicated by the observed data. For this application we assigned the misclassification costs so that misclassification by two categories was twice as costly as misclassification by one category (the costs are applied on a relative basis). Misclassification can occur due to a number of reasons including: monitoring network limitations (the highest PM concentration in an area may not be observed), use of discrete classification categories (days with PM values near the category boundaries may be misplaced into a lower or higher category, but in this case the concentration difference is only slight), the complexity of the inter-variable relationships, the completeness of the dataset with respect to defining these relationships, and data errors or missing data. The misclassification costs are used in optimizing the trees, considering both classification accuracy and the number of terminal bins.

For this study, we selected trees comprised of approximately 30 to 35 terminal bins, with the best accuracy within this size range. We examined the results with respect to classification accuracy and physical reasonableness. As discussed in Section 4 of this report, we also used the results to examine the factors influencing $PM_{2.5}$ concentrations within each area of interest. Specific review tasks included:

- The input variables and CART input parameters were checked and verified.
- The matrix representing the statistical goodness of the classification (as created by CART), was examined for serious misclassification.
- The relative importance of the input parameters was reviewed.
- The overall structure of the classification tree and number of classification bins were checked to ensure that the pathways to the different classification bins are distinct and that the bins provide a reasonable segregation of the days based on the daily $PM_{2.5}$ values.
- The values used to determine the branching of the CART output classification trees were checked to ensure that the values are reasonable and consistent with the input data.
- All splits in the decision tree were examined to ensure that the parameters and values used to develop the classification tree are physically meaningful (i.e., consistent with basic conceptual models of $PM_{2.5}$ formation and transport).
- One or more bins representing each classification category were selected and the decision pathways leading to those bins were explicitly checked to ensure that they are physically reasonable.

As a final step in the review of the CART results, we also prepared tabular summaries of the mean values of the input variables for each category and each key bin.

3.3. CART Results

As a guide to the summary of the CART results that follow, we provide a brief summary of the components of the CART application:

The CART results that are presented in the greatest detail in this section are those for which the most accurate classification was achieved for all sites, using a consistent set of data and assumptions. These CART results were used to create the final "operational" versions of the $PM_{2.5}$ forecasting tools for each area.

Prior to the preparation of the final "operational" versions of the tool for this project, draft versions of each tool were prepared and distributed to MARAMA and the state forecasters. The CART results used in the preparation of the draft tools were evaluated and refined in preparing the final results for the operational tools, and are also presented in this section.

As part of the CART application, more than 20 diagnostic and sensitivity tests were conducted for each area. The first of these included only the meteorological input parameters. The remaining tests examined the use of alternate input parameters, as well as different forms of the dependent variable for $PM_{2.5}$. The key findings from the diagnostic and sensitivity tests are discussed in this section.

As part of this exploratory analysis, a "research" version of the forecasting tool was prepared for each area that includes an additional prior day's PM concentration parameter. The "research-version" CART results are also briefly presented.

3.3.1. Summary and Key Findings from the Diagnostic and Sensitivity Testing

Meteorological and $PM_{2.5}$ Input Parameters

As first test of CART, we used only meteorological parameters. The purpose of this test was twofold, to: 1) quality assure the input datasets and ensure their readiness for CART, and 2) obtain information about the relative importance of the various meteorological parameters in the construction of the CART trees. Please note that several sensitivity tests were performed as part of the meteorological parameters only applications, to refine the meteorological inputs. A key refinement was the use of the number of six-hour periods of precipitation versus total precipitation; this was done to represent both the magnitude and the temporal (and potentially geographical) extent of the precipitation. The change in 700 mb geopotential height (from the prior day to the analysis day) replaced the twice-daily 700 mb geopotential height variables as a potentially better indicator of regional-scale pressure patterns. Both of these changes to the meteorological input parameters resulted in a slight improvement to the CART results for most areas.

We then added the prior day (two-days-ago) $PM_{2.5}$ concentrations to the input files. These results provided insight into the relative importance of the prior-day PM inputs and whether this information improved the ability of the CART tool to correctly classify the historical days and develop meaningful relationships. The results for each area follow. Please note that several sensitivity tests were conducted for each area to determine the best approach for including the prior day $PM_{2.5}$ values—these addressed which local and upwind sites/areas to include and how/whether to combine the values over multiple sites.

The results of the "met only" and combined meteorological and air quality parameters CART applications for each area are as follows:

- For Charlotte, the key meteorological parameters are the 900 mb to surface temperature difference, relative humidity, and 850 mb temperature. Surface wind speed is also a factor. Including the two-days-ago PM concentrations for Charlotte, Greenville-Spartanburg, and Winston-Salem changes the relative importance of the parameters slightly, but the meteorological parameters listed above remain most important. The overall classification accuracy of 78 percent is not improved, but the misclassification of Category 1 and 2 days into Category 3 bins is improved dramatically (reduced from 15 to 6).
- For Bristol, the key meteorological parameters are the 850 mb temperature, 900 mb to surface temperature difference, and relative humidity. Surface temperature is also somewhat important. Including the two-days-ago PM concentrations for Knoxville changes the relative importance of the parameters such that surface wind direction, surface temperature, and the two-days-ago Knoxville parameters are most important. Overall classification accuracy is improved from 83 to 88 percent, and the misclassification of Category 1 and 2 days into Category 3 bins is also improved (reduced from 11 to 4 days).
- For Roanoke, the key meteorological parameters are wind speed aloft, surface temperature, and, 850 mb temperature. Including the two-days-ago PM concentrations for Winston-Salem changes the relative importance of the parameters slightly. Overall classification accuracy is improved (from 88 to 92 percent), but the two Category 3 days are misclassified as Category 2. Including the two-days-ago information for Richmond results in almost no change to the classification tree and slight worse results; this parameter was not retained in subsequent CART applications for Roanoke.
- For Richmond, the key meteorological parameters are surface temperature and geopotential height; 900 to surface temperature difference and surface wind speed are somewhat important. Including the two-days-ago PM concentrations for Richmond, Winston-Salem, and Washington, D.C. changes the relative importance of the parameters slightly, and improves the overall classification accuracy slightly (from 83 to 84 percent). The misclassification of Category 1 and 2 days into Category 3 bins is improved (reduced from 6 to 0 days).
- For Washington, D.C., the key meteorological parameters are the 900 mb to surface temperature difference, 850 mb temperature, and surface wind speed. Surface wind direction and relative humidity area also factors. Including the two-days-ago PM concentrations for Washington, Baltimore, Gettysburg areas (the maximum over the three areas) and Richmond changes the order of importance of the key parameters. Overall classification accuracy is improved only slightly (from 76 to 77 percent), but the misclassification of Category 1 and 2 days into Category 3 bins is improved dramatically (reduced from 38 to 17 days). As we move into the Northeast Corridor, please note that there are quite a few more high $PM_{2.5}$ days.
- For Baltimore, the key meteorological parameters are the 900 mb to surface temperature difference, surface wind speed, and surface temperature. Surface wind direction and relative humidity are also important. Including the two-days-ago PM concentrations for Washington, Baltimore, Gettysburg areas (the maximum of the three areas) and Richmond changes the order of importance of the key parameters, and the PM values for the three areas moves to fourth in importance. Overall classification accuracy is improved from 74 to 80 percent and the misclassification of Category 1 and 2 days into Category 3 bins is reduced from 33 to 25 days.

- For Philadelphia, the key meteorological parameters are the surface temperature and 900 mb to surface temperature difference. Surface wind speed and relative humidity are also somewhat important. Including the two-days-ago PM concentrations for Camden and New Castle (the maximum over the two areas) and Washington, Baltimore, and Gettysburg (the maximum of the three areas) does not change the relative importance of the parameters, but the Camden-New Castle PM value takes on some importance. Overall classification accuracy is improved only slightly (from 80 to 82 percent), and the misclassification of Category 1 and 2 days into Category 3 bins is reduced from 24 to 16 days.
- For Wilmington, the key meteorological parameters are 850 mb temperature, geopotential height, and surface wind speed. Including the two-days-ago PM concentrations for Camden and New Castle (the maximum over the two areas) and Washington, Baltimore, and Gettysburg (the maximum of the three areas) does not change the relative importance of the parameters, but the Camden-New Castle PM value takes on some importance. Overall classification accuracy is unchanged from 78 percent, and the misclassification of Category 1 and 2 days into Category 3 bins is reduced from 41 to 36 days.
- For Newark, the key meteorological parameters are relative humidity, 900 mb to surface temperature difference, 850 mb temperature, and surface wind speed. Including the two-days-ago PM concentrations for Elizabeth and Camden-New Castle (the maximum over the two sites) changes the relative importance of the parameters slightly. Both PM parameters take on some importance. Overall classification accuracy is improved from 80 to 84 percent, but the misclassification of Category 1 and 2 days into Category 3 bins is worse (increased from 10 to 14 days).

The resulting CART trees using the combined meteorological and $PM_{2.5}$ parameters were designated the “Regional 1” series of trees. Key findings from the CART results at this stage of the project included:

- Different types of $PM_{2.5}$ episodes can be identified based on meteorological and prior day PM indicators.
- Regional $PM_{2.5}$ variables are more important for smaller/southern urban areas; local $PM_{2.5}$ variables are more important for the larger/more northern areas.
- Stability parameters are important for all areas.
- Temperature tends to be used as a splitter early in the tree (segregating the days seasonally).
- Relative humidity is used to segregate the days but does not have a straightforward categorical tendency.
- Wind speed is important and lower wind speeds almost always lead to higher $PM_{2.5}$ bins.
- Wind direction is often used as a split parameter, but does not always vary regularly among the categories.
- For all areas, less precipitation is associated with lower $PM_{2.5}$ but is not frequently used by CART.

Refinement of Meteorological Input Parameters

Additional sensitivity tests involved some refinement and modification of the meteorological parameters. Specifically, 700 mb wind data for the analysis day were omitted from the CART application. Those for the day prior to the analysis day were retained. Note that the 700 mb pressure level is typically at a height of approximately 3000 m. The reasoning here was that while the higher-level winds may influence the transport of pollutants into an area on the day prior, the local weather and transport conditions for the day in question are better described by the 850 mb winds. The variable use of the winds for both levels also suggested some redundancy in the information. Overall, the CART results were improved slightly when the 700 mb winds for the analysis day were omitted (mostly with regard to the reasonableness of the splits defining the pathways to the bins).

In addition, a new parameter was added to the CART analysis to indicate the time of year or season. This was primarily an attempt to represent the known variations in the amount of biogenic emissions that are present in the atmosphere and that may contribute to secondary aerosol formation. To account for seasonal variations in vegetative cover, three periods were defined. The winter period includes November, December, January, February, and March. The transitional period includes April, May, September, and October. The summer period includes June, July, and August. Including this parameter did not significantly change the CART results. Instead, surface temperature was more frequently used by CART to separate the days seasonally. Nevertheless, this parameter was retained for possible future refinement.

With these additional refinements, the resulting CART trees were designated the “Regional 2” series of trees. These were used in preparing the preliminary version of the “operational” forecasting tools.

Following an evaluation of the preliminary tools, using both real-time and historical data, additional sensitivity tests were designed and conducted to include some additional meteorological information that seemed relevant to some missed forecasts, and to incorporate some new ideas related to the use of prior day PM data.

The relative importance of stability and specifically the 900 mb to surface temperature difference parameter for most areas led us to consider whether additional stability parameters would be helpful in capturing inversions or other stability related features with different depths. In addition, for one area, Philadelphia, a missed forecast for a winter day with high observed $PM_{2.5}$ concentration seemed to be due to the presence of a very shallow surface inversion (B. Ryan, personal communication). To test the use of additional stability parameters, we defined two new parameters: the 850 mb to surface temperature difference and the 950 mb to surface temperature difference. We also defined a parameter that was the maximum of the three stability parameters, thinking that this would capture the inversion strength, regardless of the depth of the inversion. We then tested the use of these parameters in CART, first by substituting the maximum value parameter for the 900 mb difference, and then by adding all three of the difference parameters. Use of the maximum value parameter degraded the CART results for almost all areas. In hindsight, this is likely because the parameter represented different things on different days and thus it was difficult for CART to establish relationships among all days included in the dataset. Use of all three parameters, neither significantly improved nor degraded the results. CART tended to make use of all three of the parameters in various parts of the trees, and the relationships seemed reasonable. The three (separate) stability parameters were retained in subsequent of the CART applications. The resulting CART trees were designated the “Regional 3” series of trees. These were used in preparing the final version of the “operational” forecasting tools.

A second missed high $PM_{2.5}$ forecast (also for the Philadelphia area) appeared to be related to the regional-scale recirculation of pollutants (from Philadelphia to over the Atlantic Ocean, and then back again) over a three-day period (B. Ryan, personal communication). To account for this type of event we experimented with two different recirculation indexes. In both cases, the recirculation index was defined based on the 850 mb wind data already included in the CART analysis. This parameter was assigned a value of 0 or 1, with 1 indicating a potential for recirculation aloft. In the first of these, recirculation was defined according to: (1) the difference in wind direction at the 850 mb level between the previous day's evening sounding and the analysis day's morning sounding and (2) the average 850 mb wind speed (average of the evening and morning soundings). A day was classified as a recirculation day if the difference in 850 mb wind direction from the previous afternoon to the current morning was within 15 degrees of 180 degrees (i.e., almost directly opposite) or if the average wind speed at 850 mb was less than or equal to 3 ms^{-1}. In the second of these, two-day recirculation was also considered—using the same definition as above—and if either one-day or two-day recirculation was indicated, the index was set equal to 1. These parameters were included separately in CART but yielded no change in the CART results. In both cases, the index was not considered important by CART and the parameter was not retained for subsequent CART applications.

Prior-Day $PM_{2.5}$ Input Parameters

A final series of diagnostic and sensitivity simulations were conducted to examine the use of $PM_{2.5}$ data for one day prior to the analysis day (rather than two days prior). Of course, this is problematic from a forecasting perspective, since forecasts need to be made around midday and hourly data would only be available through approximately noon. There are several approaches that have been developed to estimate the air quality index using only 12 hours of hourly $PM_{2.5}$ data. Three of these are discussed and evaluated by McMillan (2004). For this study, we assumed that one or more of these approaches would be used and we included the prior day's value in CART.

For this series of tests, we prepared the prior day $PM_{2.5}$ input data three different ways, based on: 1) FRM data, 2) noon-to-noon 24-hour average of the continuous data, and 3) 12-hour average of the continuous data. These additional PM inputs were prepared for the same sites that were used to specify the two-days-ago values. In preparing the data, we found that the use of continuous data resulted in major data gaps, in many cases because the continuous monitors came on line during the mid to latter part of the analysis period. Use of these data in CART gave poor results. Instead we focused on the use of the FRM data, with the assumption that forecasters would use some methodology to estimate the prior day values.

Several alternative prior-day $PM_{2.5}$ parameters were tested. First, the prior day value was simply added to the dataset. It was used both in conjunction with the two-days-ago value, and as a replacement to the two-days-ago value. In both cases, the use of the prior-day value increased the tendency for Category 1 and 2 days to be placed into Category 3 bins. One possible explanation for this is that the meteorological parameters used in CART are not able to fully describe the conditions that would lead to a decrease in $PM_{2.5}$ (such as a cold front passage; or afternoon thundershowers). Conditions resulting in a decrease in PM are often more sudden or dramatic than those associated with an increase in PM. Thus use of a prior day value that is relatively high, frequently results (in CART) in a high value on the analysis day.

To try to mitigate the importance of the prior-day $PM_{2.5}$ concentration (as well as the need for forecasters to correctly estimate the exact value of the prior-day concentration) we also used a

binned version of the concentration as an input parameter. This new parameter was assigned a value of 1 through 4, corresponding to the following ranges in $PM_{2.5}$ concentration: less than 15.5, 15.5 to less than 28, 28 to less than 40.5, and greater than 40.5 μgm^{-3}.

Further, we calculated an adjusted prior-day $PM_{2.5}$ concentration that accounted for tendencies in the concentration. Specifically, if the difference between the prior-day and two-days-ago is positive (increasing PM concentration) no adjustment is made. If the difference between the prior-day and two-days-ago is negative (decreasing PM concentration) the prior-day value is lowered by the same percentage amount. This adjusted prior-day value was also used directly and as a binned input parameter (using the same bin structure as given above).

The results of the tests using the prior-day $PM_{2.5}$ concentrations are summarized as follows:

- Use of the prior-day $PM_{2.5}$ concentration increases the overall accuracy of the CART analysis for several areas of interest but in general the results are characterized by a greater tendency to place Category 1 and 2 days into Category 3 bins
- Binning the prior-day concentration mitigates the tendency for overestimation and lessens the importance of the parameter in the construction of the CART tree.
- Adjusting the value for decreasing PM from two-days-ago to the prior day also mitigates this tendency (by allowing for an observed decreasing tendency in $PM_{2.5}$ to be accounted for).
- Binning the adjusted prior-day concentrations gives the best results overall, for the greatest number of areas (among our areas of interest).

Other considerations also favor the binned form of the parameter. The use of a binned value alleviates the need for a forecaster to correctly estimate the value (only the range needs to be correct). Although we use three bins for the classification variable, we used four bins for the prior-day value in order to distinguish between low and high Category 2 days and account for tendencies within this rather broad category.

The resulting CART trees were designated the “Research” series of trees. These were used in preparing the “research” version of the forecasting tools.

3.3.2. Preliminary and Final “Operational” CART Results

The CART results for the preliminary and final “operational” versions of the forecasting tools are presented and compared in this section. These are designated the “Regional 2” and “Regional 3” CART trees, respectively. As noted earlier, we selected CART trees with around 30–35 bins, with the best classification accuracy possible for that range of complexity. The classification accuracy is the percentage of correctly classified days, that is, days whose concentration levels match the concentration levels of the bins in which they fall. The classification accuracy for the Regional 2 trees is 83 percent on average, ranging between 78–91 percent. For the Regional 3 trees, the average classification accuracy is 84 percent, ranging between 80 and 91 percent.

Please note that only the classification results are presented in this section of the report. A more detailed analysis of the final, operational CART results for each area is provided in Section 4.

Overall classification accuracy for the Regional 2 trees is summarized in Table 3-1. The distribution of correctly and incorrectly classified days for each classification category is provided in Table 3-2.

Table 3-1. CART Classification Matrices for Regional 2 Trees

	Number of CART Bins	Classification Accuracy (%)
Charlotte	35	82
Bristol	33	88
Roanoke	34	91
Richmond	31	83
Washington	39	78
Baltimore	33	80
Philadelphia	35	82
Wilmington	37	81
Newark	34	85

Table 3-2. CART Classification Matrices for Regional 2 Trees.

	Regional 2 CART		
Actual Class	Category 1	Category 2	Category 3
Charlotte			
1	512	93	2
2	111	434	3
3	0	0	7
Bristol			
1	187	24	0
2	17	149	4
3	0	0	8
Roanoke			
1	228	20	0
2	17	168	1
3	0	0	2
Richmond			
1	694	106	0
2	124	391	0
3	0	0	7
Washington			
1	588	139	5
2	146	474	18

	Regional 2 CART		
Actual Class	Category 1	Category 2	Category 3
3	0	2	33
Baltimore			
1	384	79	2
2	89	330	18
3	0	0	31
Philadelphia			
1	641	103	3
2	116	412	17
3	0	0	28
Wilmington			
1	562	115	2
2	106	434	18
3	0	0	26
Newark			
1	350	47	3
2	51	251	9
3	0	0	18

For Charlotte, Bristol, Roanoke, and Richmond, there are very few Category 3 days. Overall classification accuracy is good to very good, and all Category 3 days are correctly classified. There is some tendency for CART to place Category 1 and 2 days into the Category 3 bins, in particular for Charlotte and Bristol.

For the Washington, Baltimore, Philadelphia, and Wilmington areas, there are more Category 3 days and overall classification accuracy is less good, but still around 80 percent for all four areas. With the exception of two days for Washington, all of the Category 3 days are correctly classified. However, a significant number of Category 2 days (as well as some Category 1 days) are misclassified as Category 3. In general, these tend to have concentrations that are near the high end of the Category 2 range, but not in all cases. Note that the number of bins is also quite large for Washington and Wilmington, as needed to get near our target of 80 percent accuracy.

Overall classification accuracy for the Regional 3 trees is summarized in Table 3-3. The distribution of correctly and incorrectly classified days for each classification category is provided in Table 3-4.

Table 3-3. CART Classification Matrices for Regional 3 Trees

	Number of CART Bins	Classification Accuracy (%)
Charlotte	33	81
Bristol	33	90
Roanoke	34	91
Richmond	29	83
Washington	38	80
Baltimore	34	80
Philadelphia	35	82
Wilmington	36	81
Newark	34	86

Table 3-4. CART Classification Matrices for Regional 3 Trees

	Regional 3 CART		
Actual Class	Category 1	Category 2	Category 3
Charlotte			
1	486	118	3
2	91	453	4
3	0	0	7
Bristol			
1	189	21	1
2	16	151	3
3	0	0	8
Roanoke			
1	223	25	0
2	14	171	1
3	0	0	2
Richmond			
1	649	86	0
2	117	349	1
3	0	0	7

	Regional 3 CART		
Actual Class	Category 1	Category 2	Category 3
Washington			
1	596	128	8
2	141	472	25
3	0	2	33
Baltimore			
1	377	85	3
2	83	339	15
3	0	0	31
Philadelphia			
1	641	103	3
2	116	412	17
3	0	0	28
Wilmington			
1	565	110	4
2	108	437	13
3	0	0	26
Newark			
1	360	36	4
2	48	251	12
3	0	0	18

Compared to the Regional 2 trees, overall classification accuracy is about the same or slightly better and the number of bins is the same or slightly lower. The tendency for overestimation is the same for Charlotte, Bristol, Roanoke, Richmond, and Philadelphia; worse for Newark and Washington; and slightly better for Baltimore and Wilmington. In preparing the Regional 3 trees, we noted and corrected a discrepancy in our approach to omitting days from the dataset based on missing data, and for all areas consistently omitted days for which the two-days-ago PM values were missing for any of the local or upwind sites used in the CART analysis. Thus the number of days is different between the Regional 2 and Regional 3 trees for some of the areas; this is especially apparent for Richmond.

Complete listings of the CART results for the Regional 3 trees are provided as an electronic attachment to this report (Attachment B).

3.3.3. "Research" CART Results

The CART results for the "research" version of the forecasting tools are presented in this section. These are designated the "Research" CART trees. As discussed in Section 3.3.1, the Research CART trees differ from the operational CART trees in their use of prior-day $PM_{2.5}$ input parameters. The Research trees rely primarily on $PM_{2.5}$ data for one day rather than two days prior to the analysis day. The data values are adjusted using the two-day prior data to account for tendencies in the concentration and binned according to specified concentration ranges.

For the Research trees, the average classification accuracy is 84 percent, ranging between 78 and 91 percent. Although the overall accuracy is similar, these results were generally less promising than either of the "Regional" tree sets, mostly because even more days from Categories 1 and 2 were misplaced into the Category 3 bins. This is somewhat puzzling—since it makes sense that more information about prior day $PM_{2.5}$ concentrations would improve the classification rather than degrade it. This issue was not resolved as part of the current project and the research versions of the tools were developed to allow further investigation of this issue and to support future work in this area.

Overall classification accuracy for the Research trees is summarized in Table 3-5. The distribution of correctly and incorrectly classified days for each classification category is provided in Table 3-6.

Table 3-5. CART Classification Matrices for Research Trees

	Number of CART Bins	Classification Accuracy (%)
Charlotte	34	86
Bristol	29	89
Roanoke	33	91
Richmond	34	86
Washington	37	78
Baltimore	34	80
Philadelphia	33	80
Wilmington	34	82
Newark	35	87

Table 3-6. CART Classification Matrices for Research Trees

	Research CART		
Actual Class	Category 1	Category 2	Category 3
Charlotte			
1	518	75	0
2	77	446	5
3	0	0	7
Bristol			
1	182	18	4
2	14	144	7
3	0	0	8
Roanoke			
1	211	27	0
2	11	171	1
3	0	0	2
Richmond			
1	628	73	0
2	82	352	0
3	0	0	6
Washington			
1	574	148	4
2	122	478	29
3	0	0	35
Baltimore			
1	375	87	3
2	72	343	20
3	0	0	29
Philadelphia			
1	639	107	1
2	116	389	38
3	0	0	28
Wilmington			
1	537	93	5
2	92	415	20

	Research CART		
Actual Class	Category 1	Category 2	Category 3
3	0	0	24
Newark			
1	328	40	2
2	31	250	14
3	0	0	16

For all areas, all Category 3 days are correctly classified. However, a significant number of Category 2 days (as well as some Category 1 days) are misclassified as Category 3. Days for which either the prior day or two-days-ago PM values were missing for any of the local or upwind sites were omitted from the dataset. Because the criteria are applied to both prior days rather than only two-days-ago the number of is sometimes different from the Regional 3 trees.

Complete listings of the CART results for the Research trees are provided as an electronic attachment to this report (Attachment C).

This page intentionally left blank.

4. Factors Influencing $PM_{2.5}$ Concentrations

In this section, we summarize the observed data used for the CART application and use these data along with other supporting information to describe the meteorological and transport conditions associated with different $PM_{2.5}$ levels in each of the areas of interest and throughout the MARAMA region.

The description of the factors influencing $PM_{2.5}$ concentrations for each area includes 1) an analysis of the magnitude and spatial and temporal characteristics of the $PM_{2.5}$ concentrations, 2) a summary of the meteorological features influencing $PM_{2.5}$ concentrations (based on weather maps, wind distribution diagrams, local knowledge, and categorical summaries of the CART input data), and 3) a detailed analysis of the characteristics of high $PM_{2.5}$ events.

This approach to describing the $PM_{2.5}$ is designed to complement, in a qualitative sense, the forecast information provided by the CART-based $PM_{2.5}$ forecasting tool.

4.1. An Overview of the Formation, Transport, and Deposition of Fine Particulate Matter ($PM_{2.5}$) in the Atmosphere

Before we present the details of the analysis characterizing the relationships between meteorological conditions and fine particulate matter ($PM_{2.5}$) within the MARAMA region, (and the statistical analysis tool developed from this analysis to assist in forecasting $PM_{2.5}$), this section provides a brief overview of the regulatory requirements for addressing $PM_{2.5}$, and a summary of the formation, transport, and deposition processes that affect ambient concentration levels.

4.1.1. Background

As measured in the atmosphere, “fine” particles are defined as particles with diameters less than 2.5 µm, while “coarse” particles are those with observed size ranges less than 10 µm (referred to as PM_{10}). In July 1997, the National Ambient Air Quality Standard (NAAQS) for particulate matter was revised by EPA. At this time, the original annual standard for PM_{10} was retained, while a new 24-hour average PM_{10} standard was added. In addition, new annual and 24-hour $PM_{2.5}$ standards were set. As required by the Clean Air Act, EPA performed a review of the original 1997 standard in 2002 and issued a formal review in August 2003. As recommended by EPA, this review maintains the original form of the $PM_{2.5}$ standard and states that a new proposal regarding the standard will be issued in March 2005 and finalized by December 2005. As a result of these regulations, states are mandated to monitor $PM_{2.5}$ concentrations, and those that weren’t already monitoring at this time began doing so in 1999 or early 2000. To assist states in monitoring for $PM_{2.5}$, a national workshop, sponsored by EPA, was held in 1998 to address and discuss the status of PM measurement research (EPA, 1998). On the basis of data collected from 2001-2003 EPA announced, in June 2004, a list of proposed $PM_{2.5}$ nonattainment areas. Following a response by the states, final designations are expected to be provided by EPA in December 2004.

Because of the link between $PM_{2.5}$ and respiratory illness, mortality, visibility impairment, and the deposition effects on water bodies and ecosystems, much effort has been expended in recent years at both the local and national levels to assess the state of fine particle concentrations throughout the U.S., and to advance the knowledge and science of $PM_{2.5}$

formation. These efforts have been undertaken to investigate the physical and chemical processes leading to $PM_{2.5}$ formation, to establish statistical relationships between meteorology and $PM_{2.5}$ formation, and to further develop and refine existing air quality models, which will be used as planning tools to develop and evaluate control strategies for meeting the applicable standards.

In July 1999, EPA finalized a new regional haze regulation, which is aimed at protecting and improving visibility in 156 Class I areas (Wilderness Areas and National Parks). Five Regional Planning Organizations (RPOs) have been established in various parts of the country to address the requirements of the regional haze regulations. Activities being undertaken by these groups include enhanced data collection (including chemical speciation of particulate matter), data analysis, emission inventory development, and air quality modeling, which is required to show future-year improvements in visibility as a result of expected changes in precursor emissions. Using available information and the known state of the science, regional assessments have been conducted to guide certain of the RPOs' in activities aimed at addressing the regional haze rule (AER, 2001; DRI, 2002). In addition, recent reports are available that summarize the knowledge and policy implications for addressing visibility (Malm, 1999; Watson, 2002). These publications summarize the current state of knowledge and discuss the challenges to be faced in lowering $PM_{2.5}$ concentrations and improving future visibility throughout the US.

4.1.2. Formation of $PM_{2.5}$ in the Atmosphere

Fine particles (also referred to as aerosols) in the atmosphere are emitted from a variety of man-made and biogenic sources (referred to as “primary” particulates) and are formed in the atmosphere from the interaction of organic and inorganic precursor gases (referred to as “secondary” particulates). They are responsible for adverse health effects and cause the most degradation in visibility. Primary fine particulates include water droplets, dust, smoke, and soot. Emission sources include open burning, power plants, automobiles, and residential wood combustion. Secondary particulates include sulfates and nitrates, which are formed in chemical reactions of sulfur dioxide (SO_2), nitrous oxides (NO_x), reactive organic gases, ammonia, etc., which are emitted by fuel combustion sources (power plants, automobiles, heaters, boilers, etc.) and other natural sources. The chemical composition, size, and ambient levels of $PM_{2.5}$ vary widely throughout the US. Nitrates and elemental carbon make up most of the fine particle mass in the West (with sulfate a smaller constituent), while sulfate constitutes the dominant fraction in the East (followed by nitrate and carbon). Heavier particles have resident lifetimes in the atmosphere of hours (due to gravitational settling), while smaller particles have resident lifetimes of days to weeks. Smaller particles are easily inhaled into the human respiratory system and may cause physiological damage. Mercury or cadmium particles deposited out of the atmosphere are toxic to living organisms and nitrates and sulfates are corrosive to building materials and vegetation. Deposited nitrates and ammonium contribute to the eutrophication of water bodies.

The major factors that affect the concentration and distribution of $PM_{2.5}$ aerosols include:

- Spatial and temporal distribution of toxic and particulate emissions including sulfur dioxide (SO_2), oxides of nitrogen (NO_x), volatile organic compounds (VOC), and ammonia (NH_3) (both anthropogenic and nonanthropogenic),
- Size composition of the emitted PM,

- Spatial and temporal variations in the wind fields,
- Dynamics of the boundary layer, including stability and the level of mixing,
- Chemical reactions involving PM, SO_2, NO_x and other important precursor species,
- Diurnal variations of solar insulation and temperature,
- Loss of primary and secondary aerosols and toxics by dry and wet deposition, and
- Ambient air quality immediately upwind and above the region of study.

A number of reactions take place in the gas phase that lead to the formation of gases that are precursors to aerosols. Secondary aerosols are formed from gases in the atmosphere by three processes: condensation, nucleation, and coagulation. Condensation involves gases condensing on smaller nuclei, nucleation involves the interaction of gases and particles to form larger particles, and coagulation involves particle growth by collision. Relative humidity plays a key role in particle growth, especially for sulfates and nitrates.

Gaseous NO_x reacts in the atmosphere with reactive hydrocarbons and organic particulates in a very complex set of reactions resulting in secondary organic particles, nitric acid, and ammonium nitrate. Nitric acid can be a precursor to PM, but HNO_3 itself is fairly volatile and highly prone to deposition on surfaces other than PM. When ammonia is present, ammonia and nitric acid can react to form ammonium nitrate. This reaction may take place in gas phase at low humidity (forming solid particles), but it is more likely to take place in aqueous phase, in tiny water droplets (aerosols) suspended in the atmosphere. This would seem to be a straightforward process for forming PM, but the presence of sulfate (formed from SO_2, as discussed below) can cause volatile HNO3 to reform from the ammonium nitrate (Seinfeld, 1986). Therefore, the amount of PM derived from NO_x is a function not only of the rate of formation of nitric acid, but also of how much ammonia and how much sulfate is present in the atmosphere.

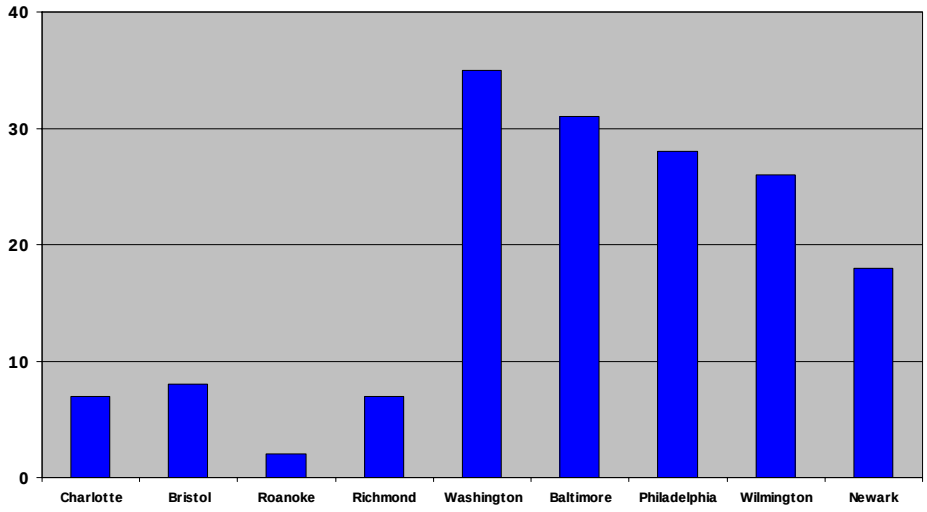
The processes leading to PM formation from SO_2 are comparatively straightforward. In the reaction with OH, SO_2 is oxidized to SO_3. When hydrated, this becomes H_2SO_4 (sulfuric acid). If ammonia gas is present, the sulfuric acid will react with it to form ammonium sulfate. Since both sulfuric acid and ammonium sulfate are strongly hygroscopic compounds, they will almost always exist in the atmosphere with a coating of water in aerosol droplets. Sulfate, therefore, exists almost exclusively in the atmosphere as PM. In the commonly used terminology, all SO_4 is referred to collectively as sulfate, whether it exists as sulfuric acid or ammonium sulfate (or other sulfate compounds).

Once fine particles are formed in the atmosphere, their small size and mass allow them to be suspended for long periods of time (days to weeks) and transported by synoptic- and meso-scale weather systems long distances from where they were originally formed. Fine dust from the Saharan Desert has been measured in the U.S., while smoke from wildfires in Central and South America and Northern Canada has also impacted areas of the U.S. It is also suspected that aerosols formed from industrial emissions in Asia travel across the Pacific to North America adding to the observed “background” aerosol concentration. Over time, depending on a number of physical factors (e.g., weather conditions, land use, etc.) fine particles deposit out of the atmosphere by both dry and wet deposition processes. Dry deposition involves settling or impaction with water bodies or other surfaces, while wet deposition includes uptake by water droplets within clouds, and subsequent rainout and washout of particles below precipitating clouds.

4.2. Regional Overview of $PM_{2.5}$

The number of days with $PM_{2.5}$ concentrations within the Unhealthy for Sensitive Groups (USG) range for the 1999–2002 CART analysis period is shown in Figure 4-1. Note that the exact number of days may be different from observed, based on our application of missing data criteria for CART, and that this chart is intended only to be used for qualitative assessment. For the four southernmost cities of Charlotte, Bristol, Roanoke, and Richmond there are a small number of USG days, ranging from 2 for Roanoke to 8 for Bristol. Keep in mind the data were only collected every third day during the analysis period for both Bristol and Roanoke, so the number of USG days for these two areas is likely somewhat higher. There is a big jump in the number of USG days as we consider the more northern sites (along the Northeast Corridor). This number drops off again further northward into New Jersey. Some missing data and different data collection start dates for the sites/areas prevent a detailed, quantitative comparison of the number of USG days, but qualitatively there seems to be a greater incidence of high $PM_{2.5}$ days in the northern part of the MARAMA study area and within the larger metropolitan areas.

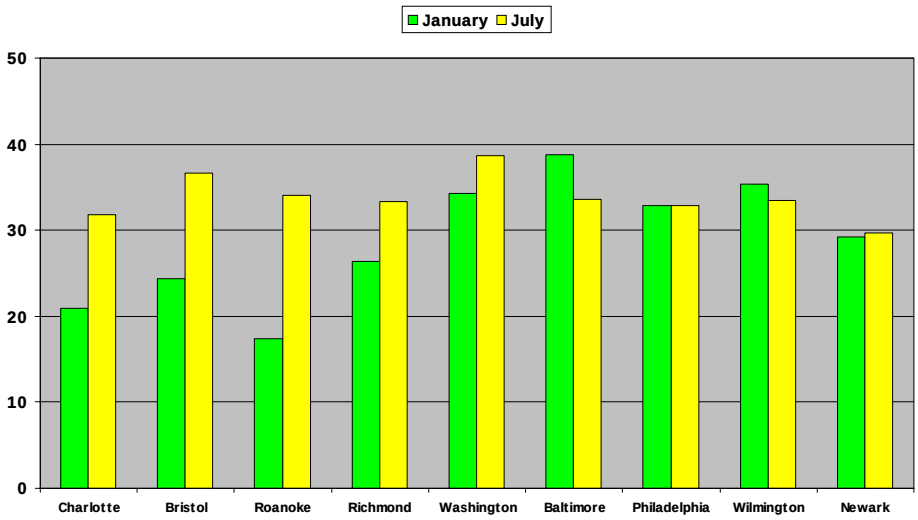
Figure 4-1. Number of USG Days During the 1999–2002 CART Analysis Period in Each of the Areas of Interest



In terms of the seasonal distribution of $PM_{2.5}$ concentrations, summer is the worst season for all nine areas and spring is almost always the best. Good concentration days are not the majority or are barely the majority for winter in Wilmington, Washington, Philadelphia, Newark, and Baltimore. Most sites see lowest PM in early spring and fall.

For the highest days (considering the 90^{th} percentile values), the $PM_{2.5}$ concentration levels are relatively consistent throughout the region during the spring, summer and fall months, but quite different during the winter months. Figure 4-2 shows the 90^{th} percentile values for January and July for each area of interest. For the more northern sites, the concentrations are higher during the winter months and the 90^{th} percentile values are consistent during these two peak periods. For the more southern sites, there are larger differences between the January and July values, with the summertime values on the order of 5 to 15 μgm^{-3} higher.

Figure 4-2. 90th Percentile Daily Maximum PM2.5 Concentration (μgm-3) for the 1999–2002 CART Analysis Period for Each of the Areas of Interest: January and July



Because determination of compliance with the $PM_{2.5}$ standard relies on quarters rather than meteorologically based seasons, it is also instructive to summarize the data in terms of the distribution of high $PM_{2.5}$ or USG days by quarter. Table 4-1 presents a general summary of the observed USG days for each of the areas of interest. Included in the table are the total number of available days of valid data contained in the datasets, and the quarterly distribution of these days. Note that the number of days given in this table is generally greater than in the final CART datasets due to missing data issues and the need to remove days with missing data for the application of CART. Because some of the PM monitors in the MARAMA area weren't deployed until 2000, and because data are available only every third day for a two of the monitors, data availability during this period varies widely from region to region, with Roanoke and Bristol having the least amount of data and the Charlotte and Washington D.C. area monitors having the most.

Table 4-1. Summary of Data Availability and Number of Observed Category 3 (Unhealthy for Sensitive Groups) Days in the MARAMA Region for the Period 1999–2002

Area of Interest	Number of days with valid PM data	Number of Observed Category 3 (USG) Days	Quarter 1 (Jan-Mar)	Quarter 2 (Apr—Jun)	Quarter 3 (Jul—Sep)	Quarter 4 (Oct—Dec)
Charlotte	1453	7	1	0	5	1
Bristol	491	7	0	2	2	3
Roanoke	469	2	0	0	2	0
Richmond	1325	7	0	0	7	0
Washington	1430	35	6	6	17	6
Baltimore	1084	31	6	10	10	5
Wilmington	1401	26	6	9	11	0
Philadelphia	1313	28	7	7	12	2
Newark	971	18	1	7	7	3

As noted in the table, the largest number of observed USG days during this period in the MARAMA region occurs either during the second and third quarters of the year, encompassing the late spring and summer periods, although some USG days occurred during the fall and winter months as well in some of the areas.

Table 4-2 summarizes the correlation between all sites, considering maximum daily $PM_{2.5}$ concentrations for all areas. R-squared values greater than 0.5 are shaded to highlight the areas of agreement. Observed concentrations for Charlotte do not appear to be well correlated with those for any of the other areas. There is some correlation between Bristol and Roanoke as well as between Roanoke and Richmond, indicating some consistency in the same-day concentrations across Virginia. Again the limited datasets for Bristol and Roanoke may limit the extent to which the R-squared values represent the similarities among these areas.

There is a slightly greater degree of correlation for Richmond and Washington, D.C. and even greater correlation for the four urban areas of Washington, Baltimore, Philadelphia, and Wilmington. The highest R-squared value is for Philadelphia and Wilmington, which are nearby to one another. There is some correlation between PM levels for Newark and those for Philadelphia, Wilmington, and to a lesser degree, Baltimore. These results suggest that there is a regional component to $PM_{2.5}$ in the areas of interest from Washington (possibly Richmond) northward, but that on any given day (with a few exceptions) there are also local meteorological and/or emissions influences that affect the areas separately. Note that these values represent same-day correlations, and do not provide the basis for discerning transport.

Table 4-2. Correlations Among the Areas of Interest: R-Squared Values Calculated Using All Daily Maximum $PM_{2.5}$ Concentrations

	Charlotte	Bristol	Roanoke	Richmond	Washington	Baltimore	Philadelphia	Wilmington	Newark
Charlotte	1.00	0.49	0.48	0.48	0.25	0.18	0.13	0.14	0.09
Bristol		1.00	0.53	0.28	0.23	0.16	0.14	0.14	0.13
Roanoke			1.00	0.53	0.45	0.36	0.28	0.32	0.23
Richmond				1.00	0.62	0.53	0.39	0.47	0.24
Washington					1.00	0.75	0.58	0.64	0.42
Baltimore						1.00	0.68	0.74	0.51
Philadelphia							1.00	0.86	0.66
Wilmington								1.00	0.58
Newark									1.00

The magnitude and distribution of $PM_{2.5}$ concentrations throughout the MARAMA region is determined in part by the prevailing meteorological conditions. Overall the location and movement of the regional-scale high- and low-pressure systems relative to an area determines the prevailing wind and dispersion conditions and thus the source-receptor relationships that characterize a $PM_{2.5}$ event, whereas the persistence and strength of the system influence/determine episode severity. A review of the meteorological conditions for days with high $PM_{2.5}$ in the areas of interest reveals that many of these days are influenced by a slow-moving or stationary high pressure system over the area of interest that results in suppressed vertical mixing of emissions/pollutants and low wind speeds or stagnation. The characteristics of high $PM_{2.5}$ events, however, vary among the areas of interest according to geographical

characteristics, local and regional emissions characteristics, and the location of each area relative to other areas in combination with pollutant-transport-conducive meteorological conditions. They also vary with season. Consequently, high $PM_{2.5}$ events occur under a variety of regional- scale and local meteorological conditions and prevailing wind directions.

In the remainder of this section, we explore the $PM_{2.5}$ concentrations and meteorological conditions influencing those concentrations for each area of interest.

4.3. Factors Influencing $PM_{2.5}$ Concentrations for Charlotte, NC

The area-wide daily maximum $PM_{2.5}$ concentration, categorized into three levels of severity, serves as the "characteristic variable" for the CART analysis and the forecasted entity for the tool. The area-wide maximum $PM_{2.5}$ for the Charlotte area was defined for this study as the maximum value over all of the sites listed as the local Charlotte sites in Table 2-1.

4.3.1. Summary of Observed $PM_{2.5}$ Data (1999–2002)

The eight FRM monitors used to determine maximum PM2.5 concentrations for the Charlotte MSA come from Cabarrus, Gaston, and Mecklenburg Counties in North Carolina, and York County in South Carolina. The monitor at the Gaston site is collocated with a second monitor, which was used to fill in data missing from the first. The dataset for Charlotte is nearly complete, (all days before August, 1999, were dropped due to missing data for a previous-day $PM_{2.5}$ variable, which was more narrowly defined). Figure 4-3 shows how days of different PM severity are distributed over the seasons. In this case the winter season is defined as December through February, spring is March through May, Summer is June through August, and Fall is September through November. "Good" days have maximum $PM_{2.5}$ concentrations less than 15.5 μgm^{-3}, "moderate" days have concentrations greater than or equal to 15.5 and less than 40.5 μgm^{-3}, and "USG" days have concentrations of 40.5 μgm^{-3} or above. USG days appear predominantly in the summer; these high-PM days are less than 1 percent of the total. Most summer days are moderate, whereas concentrations are good for most of the days in the other seasons. Figure 4-4 shows the highest 90^{th} percentile concentrations in the summer months. There are also some relatively high values in the fall (especially October) and winter months. The lowest values tend to occur during the spring.

Figure 4-3. Distribution of 1999–2002 Days by Season and Severity: Charlotte

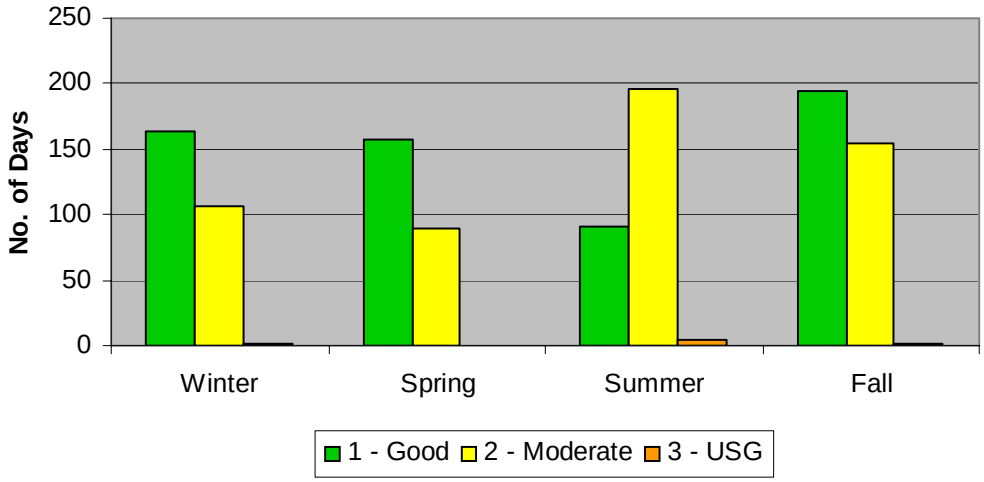
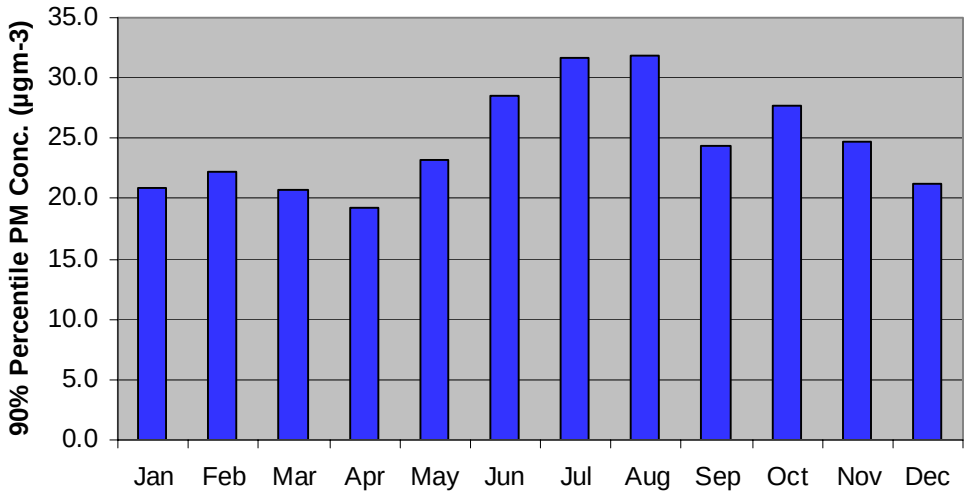


Figure 4-4. 90th Percentile Concentrations by Month (1999–2002): Charlotte



4.3.2. Meteorological Factors Influencing $PM_{2.5}$ Concentrations

The meteorological conditions associated with the different ranges of $PM_{2.5}$ concentration and specifically the highest PM days for the Charlotte area are discussed in this subsection.

Wind Patterns Associated with High $PM_{2.5}$

Plots comparing the frequency of wind directions and speeds for all low, moderate, and high $PM_{2.5}$ days for the Charlotte area are presented in Appendix A (Figure A-1). For consistency with the forecasting, low $PM_{2.5}$ days have maximum concentrations less than 15.5 μgm^{-3}, moderate days have concentrations greater than or equal to 15.5 and less than 40.5 μgm^{-3}, and high days have concentrations of 40.5 μgm^{-3} or above. The wind information in these plots is for

the Greensboro, NC upper-air monitoring site. In these diagrams, wind direction is defined as the direction from which the wind is blowing. The length of the bar within that wind-direction sector indicates the frequency of occurrence of a particular wind direction. The shading indicates the distribution of wind speeds.

Upper-air winds for the 850 mb level (approximately 1500 m above ground) are available twice per day, at approximately 0700 and 1900 EST. Distinguishing features in the wind plots (also called wind rose diagrams) for the high $PM_{2.5}$ days, when contrasted to those with other observed concentration ranges, may help to define the wind and/or transport patterns leading to high $PM_{2.5}$.

The wind roses for Charlotte are based on the Greensboro sounding data. Upper-level winds during the low PM days for Charlotte tend to be southwesterly through northwesterly for both the morning and evening soundings. Wind directions are similar for moderate PM days, with somewhat lower wind speeds, especially at the time of the evening sounding. Wind speeds are even lower for the high PM days and there is a greater tendency for easterly wind components at the time of the morning sounding.

Categorical Summaries

A comparison of the meteorological characteristics for different ranges of $PM_{2.5}$ concentration in Table 4-3 provides a basis for further distinguishing days within the different categories based on the values of meteorological parameters. In preparing this table, we used the comprehensive meteorological and PM dataset compiled for the CART application. Key meteorological parameters, as used by CART to construct the classification tree, are shaded in this table so that we can focus on the differences in these key parameters as well the differences found throughout the dataset. Categories 1 to 3 represent the standard three ranges of 24-hour $PM_{2.5}$ concentration: <15.5, 15.5-40.5, and ≥40.5 μgm^{-3}.

Table 4-3. Summary of Mean Air Quality and Meteorological Parameters for Each CART Classification Category: Charlotte

	Category 1	Category 2	Category 3
$PM_{2.5}$ Parameters			
24-hour $PM_{2.5}$ for Charlotte (μgm^{-3})	10.8	22.1	44.0
Two-days-ago 24-hour $PM_{2.5}$ for Charlotte (μgm^{-3})	13.7	16.6	23.8
Two-days-ago 24-hour $PM_{2.5}$ for Greenville-Spartanburg (μgm^{-3})	14.6	16.9	21.6
Two-days-ago 24-hour $PM_{2.5}$ for Winston-Salem (μgm^{-3})	14.2	17.0	21.3
Surface Meteorological Parameters			
Maximum surface temperature (°C)	19.4	24.2	29.8
Minimum surface temperature (°C)	8.9	11.8	16.1
Surface relative humidity (%)	66.9	67.4	65.7
Surface wind speed (ms^{-1})	2.4	1.7	1.2
Surface wind direction (degrees)	344	151	180

	Category 1	Category 2	Category 3
Number of six hour periods with precipitation (range is 1 to 4)	0.3	0.1	0.0
Upper-Air Meteorological Parameters (Greensboro)			
850 mb temperature (AM) (°C)	7.2	11.2	16.7
850 mb temperature (PM) (°C)	7.7	11.9	18.3
Temperature gradient (850 mb to surface; AM) (°C)	-2.6	-1.0	-1.9
Temperature gradient (900 mb to surface; AM) (°C)	-0.5	1.7	2.1
Temperature gradient (950 mb to surface; AM) (°C)	0.6	3.1	4.7
24-hour difference in 700 mb geopotential height (m)	-4.2	3.8	6.6
Yesterday's 700 mb wind speed (PM) (ms^{-1})	14.4	10.3	6.7
Yesterday's 850 mb wind speed (PM) (ms^{-1})	10.1	6.8	6.9
850 mb wind speed (AM) (ms^{-1})	10.9	7.3	3.4
850 mb wind speed (PM) (ms^{-1})	9.6	7.4	5.6
Yesterday's 700 mb wind direction (PM) (degrees)	269	290	333
Yesterday's 850 mb wind direction (PM) (degrees)	261	281	315
850 mb wind direction (AM) (degrees)	286	291	135
850 mb wind direction (PM) (degrees)	278	265	225
Estimated cloud cover (range of 1 to 3)	1.9	1.7	1.7
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	2	2	2

Table 4-3 provides a general overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for Charlotte. A column-by-column comparison of the values reveals some clear tendencies in several of the air quality and meteorological parameters and

High $PM_{2.5}$ in the Charlotte area is associated with relatively high $PM_{2.5}$ two-days prior—Charlotte as well as Greenville-Spartanburg and Winston-Salem. Thus, a regional day-to-day build up of $PM_{2.5}$ is indicated for high $PM_{2.5}$ days.

The surface meteorological parameters indicate a correlation between higher $PM_{2.5}$ concentrations and higher temperatures (primarily reflecting seasonal differences), lower surface wind speeds, and less precipitation. Surface wind directions tend toward southerly for the higher ranges of $PM_{2.5}$. There is no clear tendency for relative humidity.

The upper-air meteorological parameters (based here on the Greensboro sounding) indicate that higher $PM_{2.5}$ concentrations occur with higher 850 mb temperatures. There is also a tendency for more stable (positive) lapse rates to be associated with higher $PM_{2.5}$ days. This is especially true for the 900 and 950 mb temperature differences. The difference in geopotential height (defined such that a positive number indicates increasing height (pressure) over the Charlotte area) is also positively correlated with higher PM concentrations.

Lower wind speeds aloft (especially for the analysis day) and a tendency for more southerly wind directions aloft are also aligned with higher $PM_{2.5}$ concentrations.

Finally, the cloud cover and season parameters do not vary much across the three categories.

The input parameters that are most used by CART in the construction of the classification tree (either to define the splits/branching structure) or as surrogates to the primary variables in this regard are highlighted in the table. These include two-days-ago $PM_{2.5}$ for Charlotte, surface temperature, 850 mb temperature, the 950 to surface temperature difference, and 850 mb wind speed at the time of the morning sounding. All of these are also well correlated with the $PM_{2.5}$ concentration for the analysis day.

4.3.3. Characteristics of High $PM_{2.5}$ Events

The categorical summary table provides a general overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for the Charlotte area. Within the high $PM_{2.5}$ categories, there are other key differences among the parameters that result in different types of high $PM_{2.5}$ events. We have used the CART results to examine these differences.

Only certain of the CART bins are frequently associated with $PM_{2.5}$ concentrations that are in the USG range or Category 3. Of these, we identified the bins with the most number of days as key bins. Table 4-4 considers the input parameter values for the key USG bins. For Charlotte there is only one key bin and it contains four of the seven USG days.

Table 4-4. Summary of Mean Air Quality and Meteorological Parameters for the Key USG CART Classification Bin: Charlotte.

	Bin 30
Number of days	4
$PM_{2.5}$ Parameters	
24-hour $PM_{2.5}$ for Charlotte (μgm^{-3})	43.8
Two-days-ago 24-hour $PM_{2.5}$ for Charlotte (μgm^{-3})	30.9
Two-days-ago 24-hour $PM_{2.5}$ for Greenville-Spartanburg (μgm^{-3})	25.3
Two-days-ago 24-hour $PM_{2.5}$ for Winston-Salem (μgm^{-3})	26.7
Surface Meteorological Parameters	
Maximum surface temperature (°C)	35.0
Minimum surface temperature (°C)	21.3
Surface relative humidity (%)	60.2
Surface wind speed (ms^{-1})	1.3
Surface wind direction (degrees)	*
Number of six hour periods with precipitation (range is 1 to 4)	0.0

	Bin 30
Upper-Air Meteorological Parameters (Greensboro)	
850 mb temperature (AM) (°C)	18.3
850 mb temperature (PM) (°C)	20.4
Temperature gradient (850 mb to surface; AM) (°C)	-3.6
Temperature gradient (900 mb to surface; AM) (°C)	2.1
Temperature gradient (950 mb to surface; AM) (°C)	4.3
24-hour difference in 700 mb geopotential height (m)	4.3
Yesterday's 700 mb wind speed (PM) (ms^{-1})	5.3
Yesterday's 850 mb wind speed (PM) (ms^{-1})	5.6
850 mb wind speed (AM) (ms^{-1})	3.2
850 mb wind speed (PM) (ms^{-1})	5.3
Yesterday's 700 mb wind direction (PM) (degrees)	270
Yesterday's 850 mb wind direction (PM) (degrees)	225
850 mb wind direction (AM) (degrees)	180
850 mb wind direction (PM) (degrees)	198
Estimated cloud cover (range of 1 to 3)	1.8
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	3

Wind directions are evenly divided between N and S.

Since there are so few USG days, the mean characteristics of days within Bin 30 match fairly closely those for Category 3, as presented in Table 4-3 above. Even higher $PM_{2.5}$ concentrations two-days-prior, slightly lower winds speeds aloft, and a more dominant southerly wind component distinguish the Bin 30 days from the other USG/Category 3 days contained in the dataset.

While table 4-4 provides an overall summary of the mean characteristics for the key high $PM_{2.5}$ bin, it is also useful to examine the conditions associated with each day or episode.

For the Charlotte area, seven USG days occurred during the 1999–2002 period. The specific dates, including the observed $PM_{2.5}$ concentration (μgm^{-3}), are presented in Table 4-5. Included in the table is information about whether these dates are also USG days for other areas within the MARAMA region. The CART classification bin is also provided, so that the reader can link the weather summaries to the bins and characteristics discussed above.

Table 4-5. USG Days for Charlotte: 1999–2002

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	Orange Day for Other Areas
August 7, 1999	Saturday	30	41.1	
August 13, 1999	Friday	30	44.0	
January 1, 2000	Saturday	9	45.2	Washington
November 2, 2000	Thursday	28	46.9	Bristol
August 15, 2001	Wednesday	11	40.5	
July 17, 2002	Wednesday	30	45.2	
July 18, 2002	Thursday	30	44.8	Baltimore, Washington, Richmond, Newark, Philadelphia

The observed concentrations for these days generally fall in the lower end of the USG classification range. The majority of the days occur in the summer months, with two of the seven days occurring during winter.

For the August 1999 days, the Charlotte area was influenced by typical summertime surface and upper-level high-pressure systems, which affected a good portion of the Southeast. Maximum temperatures on both days approached 100°F, with mostly sunny skies and no rainfall.

On January 1, 2000, the Charlotte area was situated between a surface high-pressure system centered off the coast of Delaware and a weak low-pressure system over Mississippi. The morning lows in the area were in the upper 30's while the maximum temperatures reached the mid-60's, with light winds throughout the day and no precipitation. Since this was the first day of the new millennium, the $PM_{2.5}$ concentrations may have been influenced by early morning fireworks in the Charlotte area.

For November 2, 2000, the area was influenced by a strong upper-level ridge that affected the entire eastern seaboard. Skies were generally clear and winds were very light throughout the day, with minimum temperatures in the area in the upper 30's and maximums in the low 70's, with no precipitation.

On August 15, 2001, the weather in the Charlotte area was influenced by a weak summertime upper-level ridge and a moderately strong surface high pressure system centered over Pennsylvania. Winds were light throughout the day with minimum temperatures in the upper 60's and maximums around 90. Shallow fog conditions with 3 miles visibility were reported in the early morning hours.

For the July 17–18, 2002 period, a relatively strong upper level high was located over the southeast, with a strong surface high-pressure area over Georgia. Winds in the upper levels above Charlotte were generally very light, with a northwesterly direction. Lows during these days were near 70 with highs reaching 93 on both days. Hazy conditions were reported in the early morning on both days, with mostly sunny skies and no precipitation occurring in the area on either day. As indicated by the fact that USG days were also measured in the Baltimore, Washington, Richmond, Newark, and Philadelphia areas on July 18, the synoptic conditions causing high PM concentrations were widespread throughout the MARAMA region and persisted for several days.

This review of the meteorological conditions indicates the high PM concentrations occur under a variety of synoptic situations, but nearly all of these include high pressure over or to the north of the Charlotte area and light winds. The day-specific conditions discussed above are consistent with the categorical and CART-based average conditions for all and the subset of USG days, indicating that the CART bin captures the key characteristics of the majority of USG days and that the information contained in the categorical summaries can be used independently to guide the preparation of $PM_{2.5}$ forecasts.

4.4. Factors Influencing $PM_{2.5}$ Concentrations for Bristol, VA

The area-wide maximum $PM_{2.5}$ concentration for the Bristol area was defined for this study as the maximum value over all of the sites listed as the local Bristol sites in Table 2-1.

4.4.1. Summary of Observed $PM_{2.5}$ Data (1999–2002)

The area-wide daily maximum PM2.5 concentrations for the Bristol MSA are the daily maximums over two FRM monitors: one in Sullivan County, Tennessee, and one in the city of Bristol. A second monitor in Sullivan Co. was used as backup in the event of missing data for the first. These monitors record fine mass every three days. Two percent of the days are USG, and Figure 4-5 shows that these days only occurred in the summer and fall. Concentrations are worst in summer, which has more moderate days than good. Figure 4-6 shows the 90th percentile concentrations for each month; again, summer months have the highest value, but the November concentration follows close behind.

Figure 4-5. Distribution of 1999–2002 Days by Season and Severity: Bristol

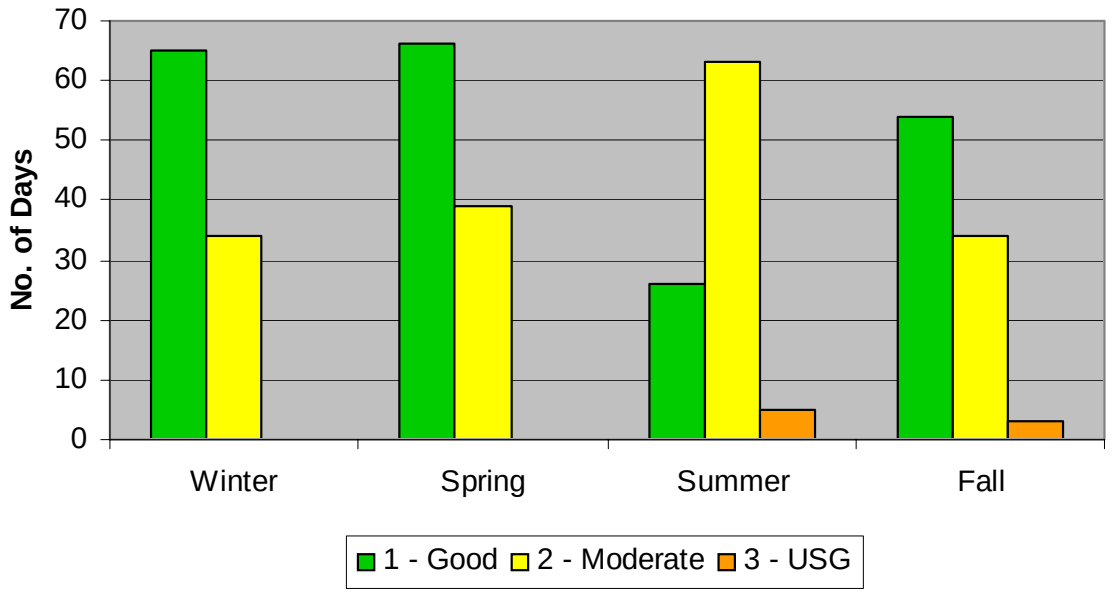
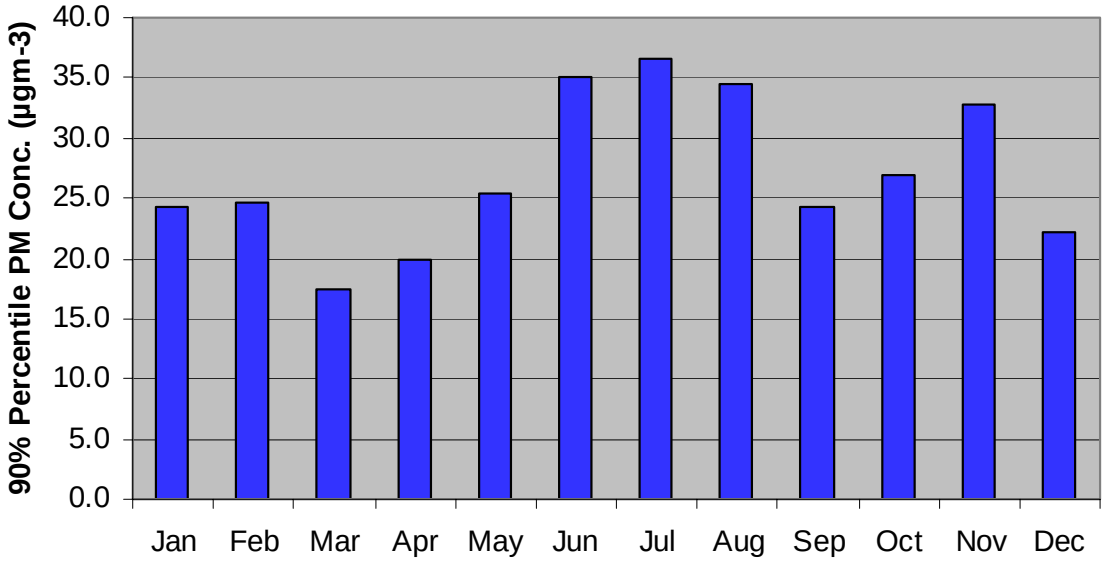


Figure 4-6. 90th Percentile Concentrations by Month (1999–2002): Bristol



4.4.2. Meteorological Factors Influencing $PM_{2.5}$ Concentrations

The meteorological conditions associated with the different ranges of $PM_{2.5}$ concentration and specifically the highest PM days for the Bristol area are discussed in this subsection.

Wind Patterns Associated with High $PM_{2.5}$

Plots comparing the frequency of wind directions and speeds for all days and for low, moderate, and high $PM_{2.5}$ days for the Bristol area are presented in Appendix A. The wind information in these plots is for the Roanoke upper-air monitoring site. The plots use the same format and contain the same information described above for the Charlotte area.

The wind roses for Bristol (Figures A-3 and A-4) are based on the Roanoke sounding data. The upper-level winds for the low PM days for Bristol tend to be westerly to northwesterly, but there are also southwesterly winds on some portion of the days. When moderate PM is observed, wind speeds are lower than for the low PM days. Compared to the low PM days, the winds are similarly directed in the morning, and there is a greater percentage of days with southwesterly winds during the evening. The highest PM days are dominated by northwesterly to northerly winds at the time of both the morning and evening sounding.

Categorical Summaries

A comparison of the meteorological characteristics for different ranges of $PM_{2.5}$ concentration in Table 4-6 provides a basis for further distinguishing days within the different categories based on the values of meteorological parameters. In preparing this table, we used the comprehensive meteorological and PM dataset compiled for the CART application. Key meteorological parameters, as used by CART to construct the classification tree, are shaded in this table so that we can focus on the differences in these key parameters as well the differences found throughout the dataset. Categories 1 to 3 represent the standard three ranges of 24-hour $PM_{2.5}$ concentration: <15.5, 15.5–40.5, and ≥40.5 μgm^{-3}.

Table 4-6. Summary of Mean Air Quality and Meteorological Parameters for Each CART Classification Category: Bristol

	Category 1	Category 2	Category 3
$PM_{2.5}$ Parameters			
24-hour $PM_{2.5}$ for Bristol (μgm^{-3})	10.4	22.5	45.0
Two-days-ago 24-hour $PM_{2.5}$ for Knoxville (μgm^{-3})	16.4	21.2	29.9
Surface Meteorological Parameters			
Maximum surface temperature (°C)	16.7	22.7	26.5
Minimum surface temperature (°C)	5.4	9.7	13.1
Surface relative humidity (%)	68.7	71.4	70.9
Surface wind speed (ms^{-1})	2.0	1.1	1.0
Surface wind direction (degrees)	278	252	270
Number of six hour periods with precipitation (range is 1 to 4)	0.4	0.2	0.1
Upper-Air Meteorological Parameters (Roanoke)			
850 mb temperature (AM) (°C)	4.1	9.9	14.0
850 mb temperature (PM) (°C)	5.1	11.2	14.7
Temperature gradient (850 mb to surface; AM) (°C)	-1.2	1.3	2.6
Temperature gradient (900 mb to surface; AM) (°C)	1.1	3.1	5.4
Temperature gradient (950 mb to surface; AM) (°C)	-7.4	3.0	-5.8
24-hour difference in 700 mb geopotential height (m)	15.6	11.8	6.8
Yesterday's 700 mb wind speed (PM) (ms^{-1})	10.9	7.5	5.6
Yesterday's 850 mb wind speed (PM) (ms^{-1})	12.0	9.7	8.9
850 mb wind speed (AM) (ms^{-1})	10.1	7.5	6.4
850 mb wind speed (PM) (ms^{-1})	274	289	270
Yesterday's 700 mb wind direction (PM) (degrees)	259	273	315
Yesterday's 850 mb wind direction (PM) (degrees)	278	278	333
850 mb wind direction (AM) (degrees)	279	263	338
850 mb wind direction (PM) (degrees)	2.0	1.9	2.0
Estimated cloud cover (range of 1 to 3)	2	2	3
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	2	2	2

A column-by-column comparison of the values in table 4-6 reveals some clear tendencies in several of the air quality and meteorological parameters.

High $PM_{2.5}$ in the Bristol area is associated with relatively high $PM_{2.5}$ in the Knoxville area two-days prior. Thus, a regional day-to-day build up is indicated.

The surface meteorological parameters indicate a correlation between higher $PM_{2.5}$ concentrations and higher temperatures, lower surface wind speeds, and less precipitation. There is no clear tendency for relative humidity or surface winds directions (which tend to be westerly, on average, for all three categories).

The upper-air meteorological parameters (based here on the Roanoke sounding) indicate that higher $PM_{2.5}$ concentrations occur with higher 850 mb temperatures. There is also a tendency for more stable (positive) lapse rates to be associated with higher $PM_{2.5}$ days. There is no clear tendency for the change in geopotential height.

Lower wind speeds aloft and a tendency for more northerly wind directions aloft are also aligned with higher $PM_{2.5}$ concentrations.

Finally, the cloud cover parameter does not vary much across the three categories, and the seasonal indicator suggests that the higher PM days tend to be during the summer months.

The input parameters that are most used by CART in the construction of the classification tree (either to define the splits/branching structure) or as surrogates to the primary variables in this regard are highlighted in the table. These include two-days-ago $PM_{2.5}$ for Knoxville, surface temperature, surface wind speed, 850 mb temperature, and 850 mb wind speed at the time of the previous evening sounding. As noted earlier, these tend also to show the greatest differences among the classification categories.

4.4.3. Characteristics of High $PM_{2.5}$ Events

The categorical summary table provides a general overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for the Bristol area. Within the high $PM_{2.5}$ categories, there are other key differences among the parameters that result in different types of high $PM_{2.5}$ events. We have used the CART results to examine these differences.

Only certain of the CART bins are frequently associated with $PM_{2.5}$ concentrations that are in the USG range or Category 3. Of these, we identified the bins with the most number of days as key bins. Table 4-7 considers the input parameter values for the key USG bins. For Bristol there is only one key bin and it contains four of the seven USG days.

Table 4-7. Summary of Mean Air Quality and Meteorological Parameters for the Key USG CART Classification Bin: Bristol

	Bin 29
Number of days	6
$PM_{2.5}$ Parameters	
24-hour $PM_{2.5}$ for Bristol (μgm^{-3})	43.2
Two-days-ago 24-hour $PM_{2.5}$ for Knoxville (μgm^{-3})	34.1
Surface Meteorological Parameters	
Maximum surface temperature (°C)	26.5
Minimum surface temperature (°C)	12.5
Surface relative humidity (%)	69.0
Surface wind speed (ms^{-1})	0.9
Surface wind direction (degrees)	270
Number of six hour periods with precipitation (range is 1 to 4)	0.2
Upper-Air Meteorological Parameters (Roanoke)	
850 mb temperature (AM) (°C)	14.4
850 mb temperature (PM) (°C)	14.8
Temperature gradient (850 mb to surface; AM) (°C)	4.4
Temperature gradient (900 mb to surface; AM) (°C)	7.2
Temperature gradient (950 mb to surface; AM) (°C)	na
24-hour difference in 700 mb geopotential height (m)	-11.9
Yesterday's 700 mb wind speed (PM) (ms^{-1})	6.1
Yesterday's 850 mb wind speed (PM) (ms^{-1})	5.2
850 mb wind speed (AM) (ms^{-1})	9.8
850 mb wind speed (PM) (ms^{-1})	7.2
Yesterday's 700 mb wind direction (PM) (degrees)	315
Yesterday's 850 mb wind direction (PM) (degrees)	315
850 mb wind direction (AM) (degrees)	333
850 mb wind direction (PM) (degrees)	326
Estimated cloud cover (range of 1 to 3)	1.8
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	2

Since there are so few USG days, the mean characteristics of days within Bin 29 match very closely those for Category 3, as presented in Table 4-6 above. Even higher $PM_{2.5}$ concentrations two-days-prior distinguish the Bin 29 days from the other USG/Category 3 days contained in the dataset. These days also tend to occur during the transitional seasons, rather than in summer.

It is also useful to examine the conditions associated with each day or episode.

For the Bristol area, seven orange days occurred during the 1999–2002 period. The specific dates, including the observed $PM_{2.5}$ concentration ($\mu g/m^3$), are listed in Table 4-8.

Table 4-8. USG Days for Bristol: 1999–2002

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG Day for Other Areas?
June 2, 2000	Friday	29	40.8	Baltimore
June 11, 2000	Sunday	29	42.2	Baltimore, Wilmington, Philadelphia, Newark
July 26, 2000	Wednesday	28	56.2	
October 24, 2000	Tuesday	29	43.0	
October 27, 2000	Friday	29	43.8	Washington, Newark
November 2, 2000	Thursday	29	43.6	Charlotte
July 18, 2001	Wednesday	29	45.7	

The observed concentrations for these days generally fall in the lower end of the USG classification range, with the exception of July 26. The USG days are distributed over the summer and fall months, with three of the eight days occurring during the fall.

For June 2, 2000, the area was influenced by a weak upper-level ridge and a surface high centered over Tennessee. Minimum temperatures were in the upper 60's, with highs in the upper 80's to low 90's. Upper-level winds were very light and westerly and no precipitation was reported in the area.

For June 11, 2000, Bristol's weather was dominated by a weak ridge aloft and a strong surface Bermuda high-pressure system centered offshore of North Carolina. Winds were light and variable on the surface throughout the day and light and southerly aloft. Shallow fog was reported in the area in the early morning hours, with lows in the upper 60's and high near 90, and no precipitation. PM concentrations were also measured in the USG range at sites in Baltimore, Wilmington, Philadelphia, and Newark, reflecting region-wide stagnation conditions across the area.

On July 26, 2000, the southeast was under the influence of a very weak upper-level ridge system, with very light winds. A surface low-pressure system was located over the Baltimore-Washington area, but hazy skies and light winds persisted in the Bristol area. Maximum temperatures were in the mid-80's, with minimums in the upper 60's. No precipitation was reported in the general area on this day.

For the October 24 and 27, 2000 days, the Bristol area was under the influence of a relatively strong upper-level ridge and strong surface high-pressure system centered over the southeast. Winds aloft on these days were very light and northwesterly. Lows were in the mid-50s and highs were near 80, with shallow fog reported both mornings and no precipitation reported either day.

For November 2, 2000, the area was influenced by a strong upper-level ridge that influenced the entire eastern seaboard. Skies were generally clear and winds were very light throughout the day, with minimum temperatures in the area in the upper 30's and maximums in the low 70's, with no precipitation. This day was also a USG day for the Charlotte area.

For July 18, 2001, Bristol's weather was influenced by a strong upper-level ridge centered over the mid-plains, and a moderately strong surface high-pressure system over Georgia. Lows were near 70 and highs approached 90 throughout the area. Winds aloft were very light and northwesterly. Hazy conditions and limited visibility were reported during the morning hours and precipitation was reported in the Roanoke area, northeast of Bristol.

This review of the meteorological conditions indicates the high PM concentration occur under a variety of synoptic situations, but nearly all of these manifest themselves as stagnation conditions near the surface. This is consistent with the very light wind speeds indicated by the categorical and CART-based average conditions for all and the subset of USG days. CART finds this parameter to be important and thus appears to capture the key characteristics of the majority of USG days. This consistency also suggests that the categorical summaries for Bristol can be used independently to guide the preparation of $PM_{2.5}$ forecasts.

4.5. Factors Influencing $PM_{2.5}$ Concentrations for Roanoke, VA

The area-wide maximum $PM_{2.5}$ for the Roanoke area was defined for this study as the maximum value over all of the sites listed as the local Roanoke sites in Table 2-1.

4.5.1. Summary of Observed $PM_{2.5}$ Data (1999–2002)

The 436 days for the Roanoke daily maximum PM concentrations come from the maximum of two FRM monitors, one in the city of Roanoke and the other in the city of Salem, Virginia. Half a percent of these days are USG, and these all occur in the summer, as Figure 4-7 shows. Most summer days are moderate and most days in the other seasons are good; the profile of monthly 90^{th} percentile concentrations shown in Figure 4-8 peaks relatively gently in July, with a minor peak in February.

Figure 4-7. Distribution of 1999–2002 Days by Season and Severity: Roanoke

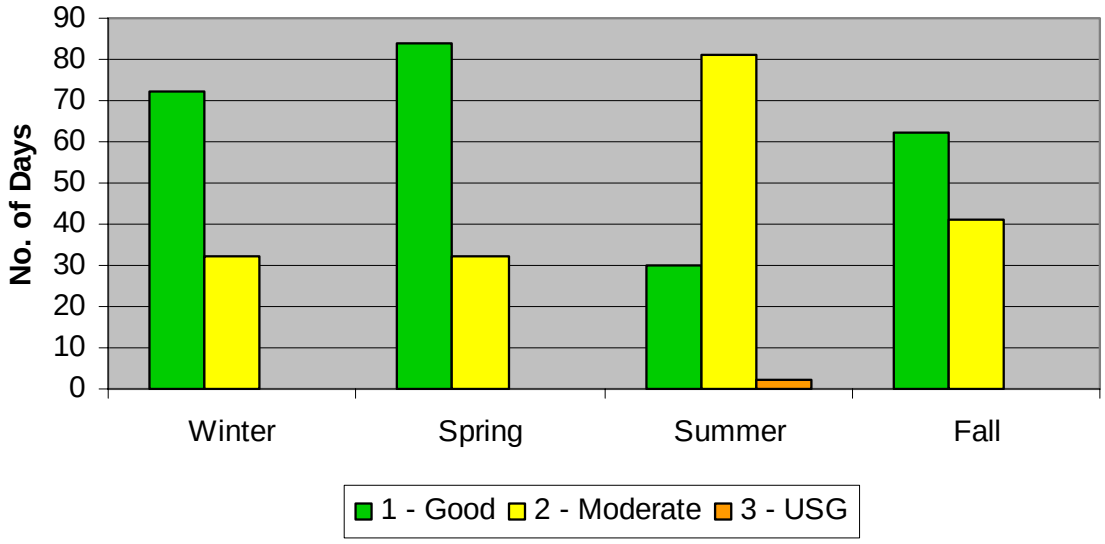
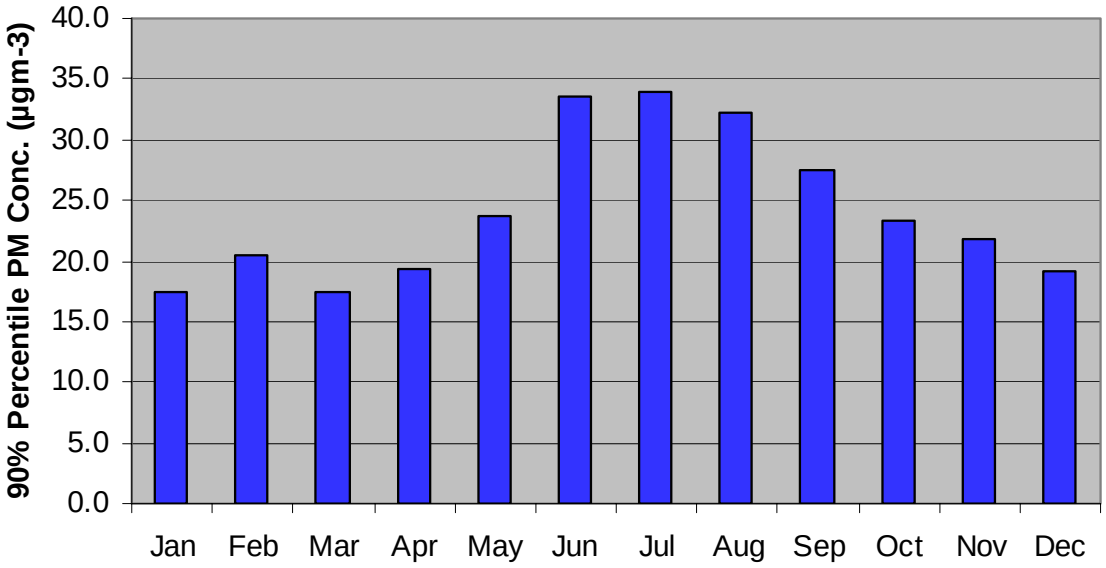


Figure 4-8. 90th Percentile Concentrations by Month (1999–2002): Roanoke



4.5.2. Meteorological Factors Influencing $PM_{2.5}$ Concentrations

The meteorological conditions associated with the different ranges of $PM_{2.5}$ concentration and specifically the highest PM days for the Roanoke area are discussed in this subsection.

Wind Patterns Associated with High $PM_{2.5}$

Plots comparing the frequency of wind directions and speeds for all days and for low, moderate, and high $PM_{2.5}$ days for the Roanoke area are presented in Appendix A. The wind information in these plots is for the Roanoke upper-air monitoring site. The plots use the same format and contain the same information as for the other areas (described earlier in this section).

The wind roses for Roanoke (Figures A-5 and A-6) are based on the Roanoke sounding data. The upper-level winds are predominately westerly to northwesterly for the low PM days. There is a notable increase in the incidence of southwesterly winds for the moderate PM days. At the time of the evening sounding, southwesterly winds dominate the wind rose. Since there are only two high PM days for Roanoke, wind roses were not prepared for this concentration level.

Categorical Summaries

A comparison of the meteorological characteristics for different ranges of $PM_{2.5}$ concentration in Table 4-9 provides a basis for further distinguishing days within the different categories based on the values of meteorological parameters. In preparing this table, we used the comprehensive meteorological and PM dataset compiled for the CART application. Key meteorological parameters, as used by CART to construct the classification tree, are shaded in this table so that we can focus on the differences in these key parameters as well the differences found throughout the dataset. Categories 1 to 3 represent the standard three ranges of 24-hour $PM_{2.5}$ concentration: <15.5, 15.5–40.5, and ≥40.5 μgm^{-3}.

Table 4-9. Summary of Mean Air Quality and Meteorological Parameters for Each CART Classification Category: Roanoke

	Category 1	Category 2	Category 3
$PM_{2.5}$ Parameters			
24-hour $PM_{2.5}$ for Roanoke (μgm^{-3})	10.1	22.7	46.8
Two-days-ago 24-hour $PM_{2.5}$ for Winston-Salem (μgm^{-3})	13.7	18.2	20.9
Surface Meteorological Parameters			
Maximum surface temperature (°C)	16.4	23.9	33.6
Minimum surface temperature (°C)	6.2	12.4	21.4
Surface relative humidity (%)	59.3	64.9	62.2
Surface wind speed (ms^{-1})	3.0	1.8	1.4
Surface wind direction (degrees)	277	230	315
Number of six hour periods with precipitation (range is 1 to 4)	0.3	0.2	0.0
Upper-Air Meteorological Parameters (Roanoke)			
850 mb temperature (AM) (°C)	4.0	11.4	18.6
850 mb temperature (PM) (°C)	5.0	12.8	20.0
Temperature gradient (850 mb to surface; AM) (°C)	-0.9	1.7	0.4
Temperature gradient (900 mb to surface; AM) (°C)	1.1	3.7	4.0
Temperature gradient (950 mb to surface; AM) (°C)	na	na	na
24-hour difference in 700 mb geopotential height (m)	-3.7	3.7	-11.5
Yesterday's 700 mb wind speed (PM) (ms^{-1})	15.8	10.1	4.4
Yesterday's 850 mb wind speed (PM) (ms^{-1})	10.9	6.6	4.1
850 mb wind speed (AM) (ms^{-1})	12.2	9.0	10.3
850 mb wind speed (PM) (ms^{-1})	10.0	7.3	3.1
Yesterday's 700 mb wind direction (PM) (degrees)	278	287	0
Yesterday's 850 mb wind direction (PM) (degrees)	270	269	0
850 mb wind direction (AM) (degrees)	284	275	0
850 mb wind direction (PM) (degrees)	277	263	270
Estimated cloud cover (range of 1 to 3)	2.0	1.8	1.0
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	2	2	3

Table 4-9 provides an overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for Roanoke. A column-by-column comparison of the values reveals some clear tendencies in several of the air quality and meteorological parameters.

High $PM_{2.5}$ in the Roanoke area is associated with relatively high $PM_{2.5}$ two-days prior—in Roanoke and to a lesser extent Winston-Salem. Thus, the regional day-to-day build up of $PM_{2.5}$ is indicated for high $PM_{2.5}$ days, with emphasis on a local build up or recirculation.

The surface meteorological parameters indicate a correlation between higher $PM_{2.5}$ concentrations and higher temperatures, lower surface wind speeds, and less precipitation. Surface wind directions tend toward northwesterly for the higher ranges of $PM_{2.5}$. There is no clear tendency for relative humidity.

The upper-air meteorological parameters (based here on the Roanoke sounding) indicate that higher $PM_{2.5}$ concentrations occur with higher 850 mb temperatures. There is also a tendency for more stable (positive) lapse rates to be associated with higher $PM_{2.5}$ days. There is no clear tendency for the difference in geopotential height.

Lower wind speeds aloft (with the exception of the 850 mb winds for the morning of the analysis day) are aligned with higher $PM_{2.5}$ concentrations. Wind directions veer from westerly to northerly with the higher PM values.

Finally, cloud cover is less for the high PM days, and the season index indicates that the highest concentrations tend to occur during the summer months.

The input parameters that are most used by CART in the construction of the classification tree (either to define the splits/branching structure) or as surrogates to the primary variables in this regard are highlighted in the table. For Roanoke, the most important parameters are surface temperature and 850 mb temperature. Wind speeds aloft are next most important. All of these are also very well correlated with the $PM_{2.5}$ concentration for the analysis day.

4.5.3. Characteristics of High $PM_{2.5}$ Events

The categorical summary table provides a general overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for the Roanoke area. Since there are only two high $PM_{2.5}$ days in the dataset for Roanoke, we did not prepare a separate table of the characteristics of the USG bins for this area.

Only two USG days occurred during the 1999–2002 period in the Roanoke area, although as noted above, the available data are limited for this site. The specific dates, including the observed $PM_{2.5}$ concentration ($\mu g/m^3$), are as follows:

Table 4-10. USG Days for Roanoke: 1999–2002

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG Day for Other Areas?
August 8, 2001	Wednesday	32	52.8	Baltimore, Washington, Richmond
July 1, 2002	Monday	33	40.7	Baltimore, Washington (7/2)

Both of these USG days occur in the summer with observed concentrations for August 8, 2001 in the middle of the USG range, while the concentration for July 1, 2002 is just within the USG category.

On August 8, 2001, the Roanoke area was influenced by a broad upper-level ridge centered over the central U.S., and a weak surface high pressure system over Georgia. Surface winds throughout the day were calm with upper-level winds very light and southerly. The low for Roanoke was 74°F while the high for this day was 93, with no precipitation reported. Similar conditions persisted throughout the region leading to high PM concentrations in the Baltimore, Washington, and Richmond areas on this day.

For July 1, 2002, the Roanoke weather was dominated by a strong upper-level ridge centered over the Midwest, and a strong, broad surface high pressure system centered directly over the Roanoke area. Upper-level winds on this day were very light and variable, while surface winds were light and variable. The low for Roanoke was 68, while the high for the day was 88. This day was also a USG day for the Baltimore and Washington areas and was the start of the multi-day PM episode across the MARAMA region which lasted through July 4th, as the upper-level ridge built further over the area, strengthening the persistent surface high. USG days occurred in the Baltimore, Washington, Philadelphia, and Newark areas on July 2 and 3, and in Washington on July 4.

This review of the meteorological conditions indicates the high PM concentration occurs in conjunction with surface high pressure and light winds, allowing for the multi-day build up of particulates in the area. This is consistent with the very light wind speeds indicated by the categorical averages. CART appears to capture the effects of the high pressure using the 850 mb temperature as a key parameter in distinguishing the high PM days. There are really not enough high PM days for Roanoke to say much more about the characteristics of the high PM days.

4.6. Factors Influencing $PM_{2.5}$ Concentrations for Richmond, VA

The area-wide maximum $PM_{2.5}$ for the Richmond area was defined for this study as the maximum value over all of the sites listed as the local Richmond sites in Table 2-1.

4.6.1. Summary of Observed $PM_{2.5}$ Data (1999–2002)

Maximum PM2.5 concentrations over five FRM monitors in Charles City, Richmond City, Chesterfield County, and Henrico County determined the area-wide maximum for Richmond. Of the days with available data from the 1999–2002 period, about half a percent had USG concentrations, and all of these occurred in summer, as shown in Figure 4-9. The majority of summer days were moderate, whereas good days dominated the other seasons. In Figure 4-10, which shows the 90th percentile concentrations, one sees highest concentrations in the summer months, and the next highest in January. Richmond is characterized by a more distinct annual profile than many of the other areas included in the analysis.

Figure 4-9. Distribution of 1999–2002 Days by Season and Severity: Richmond

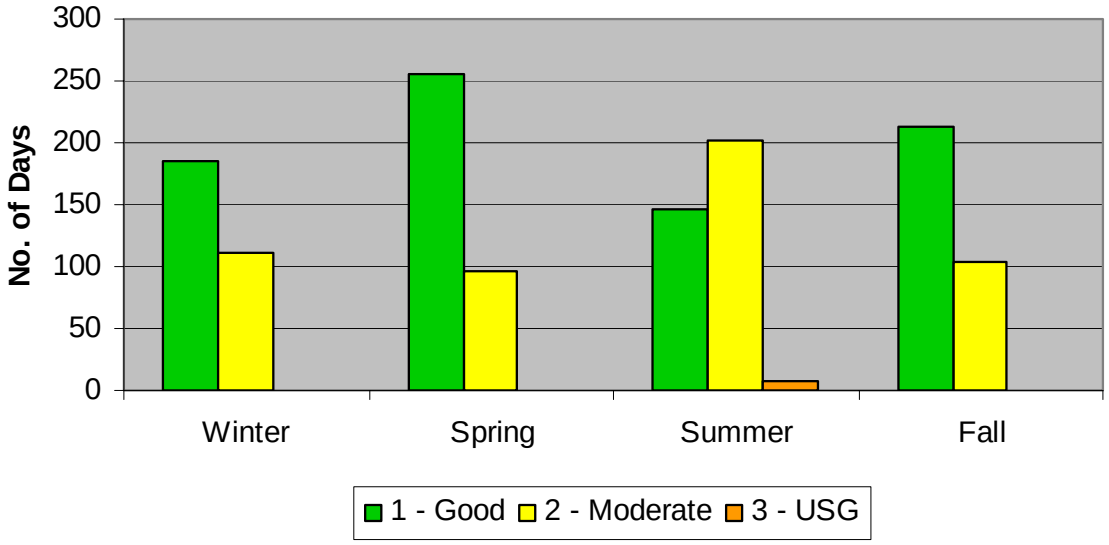
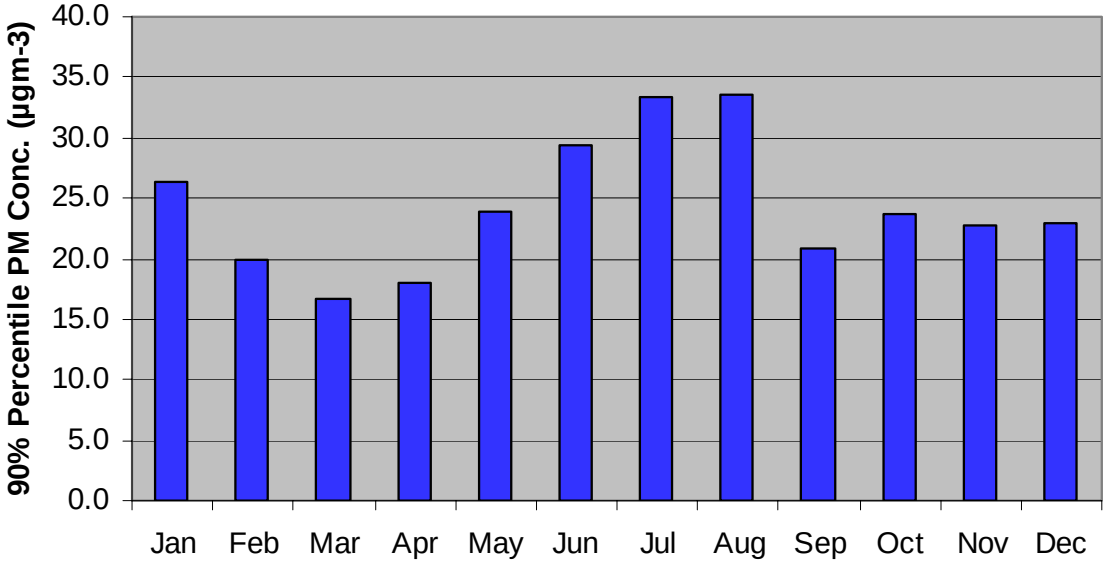


Figure 4-10. 90th Percentile Concentrations by Month (1999–2002): Richmond



4.6.2. Meteorological Factors Influencing $PM_{2.5}$ Concentrations

The meteorological conditions associated with the different ranges of $PM_{2.5}$ concentration and specifically the highest PM days for the Richmond area are discussed in this subsection.

Wind Patterns Associated with High $PM_{2.5}$

Plots comparing the frequency of wind directions and speeds for all days and for low, moderate, and high $PM_{2.5}$ days for the Richmond area are presented in Appendix A. The wind information in

these plots is for the Dulles Airport (Sterling, VA) upper-air monitoring site. The plots use the same format and contain the same information as for the other areas (described earlier in this section).

The wind roses for Richmond (Figures A-7 and A-8) are based on the Dulles Airport sounding data. The upper-level winds are predominately southwesterly to northerly for the low PM days, for both the morning and evening sounding. For many of the days, the directions fall within the westerly to northwesterly portion of this range. There is a notable increase in the incidence of southwesterly winds for the moderate PM days; wind speeds are also lower for the moderate days. At the time of the evening sounding, southwesterly winds dominate the wind rose. For the highest PM days, winds have either a northerly or southerly component. Given the small number of days, a wind pattern does not emerge. Wind speeds are much lower than for the other PM concentration levels.

Categorical Summaries

A comparison of the meteorological characteristics for different ranges of $PM_{2.5}$ concentration in Table 4-11 provides a basis for further distinguishing days within the different categories based on the values of meteorological parameters. In preparing this table, we used the comprehensive meteorological and PM dataset compiled for the CART application. Key meteorological parameters, as used by CART to construct the classification tree, are shaded in this table so that we can focus on the differences in these key parameters as well the differences found throughout the dataset. Categories 1 to 3 represent the standard three ranges of 24-hour $PM_{2.5}$ concentration: <15.5, 15.5–40.5, and ≥40.5 μgm^{-3}.

Table 4-11. Summary of Mean Air Quality and Meteorological Parameters for Each CART Classification Category: Richmond

	Category 1	Category 2	Category 3
$PM_{2.5}$ Parameters			
24-hour $PM_{2.5}$ for Richmond (μgm^{-3})	10.1	22.1	46.1
Two-days-ago 24-hour $PM_{2.5}$ for Richmond (μgm^{-3})	13.5	16.2	25.3
Two-days-ago 24-hour $PM_{2.5}$ for Washington, D.C. (μgm^{-3})	15.0	18.1	31.2
Two-days-ago 24-hour $PM_{2.5}$ for Winston-Salem (μgm^{-3})	14.5	17.6	27.5
Surface Meteorological Parameters			
Maximum surface temperature (°C)	18.7	23.8	35.8
Minimum surface temperature (°C)	8.2	12.0	23.2
Surface relative humidity (%)	67.2	70.4	63.6
Surface wind speed (ms^{-1})	3.1	2.3	1.8
Surface wind direction (degrees)	318	188	171
Number of six hour periods with precipitation (range is 1 to 4)	0.3	0.2	0.1

	Category 1	Category 2	Category 3
Upper-Air Meteorological Parameters (Dulles Airport)			
850 mb temperature (AM) (°C)	4.5	9.9	18.8
850 mb temperature (PM) (°C)	5.3	10.6	19.6
Temperature gradient (850 mb to surface; AM) (°C)	-2.9	-1.0	-3.0
Temperature gradient (900 mb to surface; AM) (°C)	-0.5	2.1	2.1
Temperature gradient (950 mb to surface; AM) (°C)	0.2	3.1	3.6
24-hour difference in 700 mb geopotential height (m)	-3.9	2.6	-18.4
Yesterday's 700 mb wind speed (PM) (ms^{-1})	11.3	8.3	4.5
Yesterday's 850 mb wind speed (PM) (ms^{-1})	10.6	8.7	5.1
850 mb wind speed (AM) (ms^{-1})	14.6	11.7	5.2
850 mb wind speed (PM) (ms^{-1})	10.9	8.2	6.7
Yesterday's 700 mb wind direction (PM) (degrees)	293	280	315
Yesterday's 850 mb wind direction (PM) (degrees)	283	269	270
850 mb wind direction (AM) (degrees)	281	286	315
850 mb wind direction (PM) (degrees)	274	281	270
Estimated cloud cover (range of 1 to 3)	1.9	1.8	1.7
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	2	2	3

Table 4-11 shows how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for Richmond. A column-by-column comparison of the values reveals some clear tendencies in several of the air quality and meteorological parameters.

High $PM_{2.5}$ in the Richmond area is associated with relatively high $PM_{2.5}$ two-days prior—in Richmond, Washington, D.C., and Winston-Salem. Thus, a regional day-to-day build up of $PM_{2.5}$ is indicated for high PM days.

The surface meteorological parameters indicate that higher $PM_{2.5}$ concentrations occur with higher temperatures (primarily reflecting seasonal differences), lower surface wind speeds, lower relative humidity, and less precipitation. Surface wind directions are northwesterly, on average for the low PM days, and tend toward southerly for the higher ranges of $PM_{2.5}$.

The upper-air meteorological parameters (based here on the Dulles Airport sounding) indicate that higher $PM_{2.5}$ concentrations occur with higher 850 mb temperatures and somewhat greater stability (although the relationships between PM and stability is less well defined than for some of the other areas, possibly due to distance and location of the upper-air monitoring site). There is no clear trend in the difference in geopotential height parameter.

There is a very clear tendency for lower wind speeds aloft (for both the day prior to the analysis day and the analysis day) but little difference in wind directions aloft among the categories.

The cloud cover parameters do not vary much, and the seasonal index show that most of the USG days occur during summer.

The input parameters that are most used by CART in the construction of the classification tree (either to define the splits/branching structure) or as surrogates to the primary variables in this regard are highlighted in the table. These include two-days-ago $PM_{2.5}$ for Winston-Salem, surface temperature, and 850 mb temperature. Surface wind speed is also somewhat important. The upper-level wind speeds appear to vary directly with PM, but are of lesser importance in the construction of the CART tree.

4.6.3. Characteristics of High $PM_{2.5}$ Events

The categorical summary table provides a general overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for the Richmond area. Within the high $PM_{2.5}$ categories, there are other key differences among the parameters that result in different types of high $PM_{2.5}$ events. We have used the CART results to examine these differences.

Only certain of the CART bins are frequently associated with $PM_{2.5}$ concentrations that are in the USG range or Category 3. Of these, we identified the bins with the most number of days as key bins. Table 4-12 considers the input parameter values for the key USG bins. For Richmond there is only one key bin and it contains all of the seven USG days.

Table 4-12. Summary of Mean Air Quality and Meteorological Parameters for Key USG CART Classification Bins: Richmond

	Bin 27
Number of days	7
$PM_{2.5}$ Parameters	
24-hour $PM_{2.5}$ for Richmond (μgm^{-3})	46.1
Two-days-ago 24-hour $PM_{2.5}$ for Richmond (μgm^{-3})	25.3
Two-days-ago 24-hour $PM_{2.5}$ for Washington, D.C. (μgm^{-3})	31.2
Two-days-ago 24-hour $PM_{2.5}$ for Winston-Salem (μgm^{-3})	27.5
Surface Meteorological Parameters	
Maximum surface temperature (°C)	35.8
Minimum surface temperature (°C)	23.2
Surface relative humidity (%)	63.6
Surface wind speed (ms^{-1})	1.8
Surface wind direction (degrees)	171

	Bin 27
Number of six hour periods with precipitation (range is 1 to 4)	0.1
Upper-Air Meteorological Parameters (Dulles Airport)	
850 mb temperature (AM) (°C)	18.8
850 mb temperature (PM) (°C)	19.6
Temperature gradient (850 mb to surface; AM) (°C)	-3.0
Temperature gradient (900 mb to surface; AM) (°C)	2.1
Temperature gradient (950 mb to surface; AM) (°C)	3.6
24-hour difference in 700 mb geopotential height (m)	-18.4
Yesterday's 700 mb wind speed (PM) (ms^{-1})	4.5
Yesterday's 850 mb wind speed (PM) (ms^{-1})	5.1
850 mb wind speed (AM) (ms^{-1})	5.2
850 mb wind speed (PM) (ms^{-1})	6.7
Yesterday's 700 mb wind direction (PM) (degrees)	315
Yesterday's 850 mb wind direction (PM) (degrees)	270
850 mb wind direction (AM) (degrees)	315
850 mb wind direction (PM) (degrees)	270
Estimated cloud cover (range of 1 to 3)	1.7
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	3

Since the seven USG days are all contained in Bin 27, the characteristics of this bin are identical to those for the Category 3 days, as discussed above

Next we examine the conditions associated with each high PM day.

Data retrieval for the Richmond area was high for the period 1999–2002, and only seven USG days occurred during this period. The specific dates, including the observed $PM_{2.5}$ concentration (μgm^{-3}), are listed in Table 4-13.

Table 4-13. USG days for Richmond: 1999–2002

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG Day for Other Areas?
July 6, 1999	Tuesday	27	48.5	
August 8, 2001	Wednesday	27	51.5	Roanoke, Baltimore, Washington
August 9, 2001	Thursday	27	41.5	Baltimore, Washington, Philadelphia, Wilmington
July 2, 2002	Tuesday	27	41.6	Baltimore, Washington, Philadelphia, Newark
July 3, 2002	Wednesday	27	50.5	Baltimore, Washington, Philadelphia, Newark
July 18, 2002	Thursday	27	46.2	Baltimore, Washington, Charlotte, Newark
August 13, 2002	Tuesday	27	42.2	Baltimore, Washington, Wilmington, Philadelphia

All of the USG days for the Richmond area occurred during the summer months, with observed concentrations in the middle of the range for the USG category. Except for the July 6, 1999 period, the meteorological conditions in the MARAMA region were widespread and persistent enough to cause high PM concentrations throughout the entire domain on the USG days measured in the Richmond area.

For July 6, 1999, the weather in the Richmond area was influenced by a broad, relatively flat upper-level ridge, and by a surface high-pressure system centered over Mississippi. Winds aloft were weak and southwesterly while surface winds were light and variable. The low temperature for the day at Richmond was 75, while the high was 98. Hazy skies and limited visibility were reported in the early morning hours and no precipitation occurred in the area on this day.

For the August 8–9, 2001 period, the Richmond area and the State of Virginia were influenced by a broad upper-level ridge centered over the central U.S., and a weak surface high-pressure system over Georgia. Surface winds throughout the day were calm with upper-level winds very light and southerly. The low temperatures on these two days in Richmond were in the low 70's, while the highs were in the upper 90's, with no precipitation reported in the area on either of these days. August 8, 2001 was also a USG day for the Roanoke area, so conditions conducive to the buildup of PM were pervasive across the state.

As noted above, the July 1–4, 2002 period exhibited high PM conducive conditions throughout the MARAMA region, with USG days measured from the Richmond area and at all sites north during this multi-day episode. Conditions are discussed in an earlier section.

The July 18–19, 2002 period exhibited severe, PM conducive conditions during which USG days were measured throughout the MARAMA region. The region was under the influence of a broad summertime upper-level ridging pattern that was transitioning to weak zonal flow. A surface high-pressure system centered over Georgia resulted in stagnant winds, high temperatures, and mostly clear, hazy skies throughout the region. High temperatures were in the mid-90s, while lows were measured in the low 70's. Hazy skies and limited visibility were reported in the Richmond area during this period.

This review of the meteorological conditions indicates the high PM concentration occur when the region is under the influence of a high-pressure system; conditions near the surface are characterized by high temperatures and low wind speeds. These conditions are consistent with categorical and CART-based average conditions for the USG days. CART primarily uses surface temperature, 850 mb temperature, and surface wind speed to represent these conditions. CART also picks up on the regional-scale build up or PM as a precursor of USG days.

4.7. Factors Influencing $PM_{2.5}$ Concentrations for Washington, D.C.

The area-wide maximum $PM_{2.5}$ for the Washington area was defined for this study as the maximum value over all of the sites listed as the local Washington sites in Table 2-1.

4.7.1. Summary of Observed $PM_{2.5}$ Data (1999–2002)

Eleven FRM monitors, plus two additional monitors used to fill in missing data points for their respective collocated monitors, determined the area maximum for Washington DC. Of the days examined, 2.5% are USG, and these are spread over all seasons, with half occurring in summer, winter and fall each taking about a quarter, and one lone high PM day appearing in the spring. Figure 4-11 visualizes this distribution, and also shows closely matched quantities of good and moderate days in the winter, a prevalence of good days in the spring, mostly moderate days in the summer, and mostly good days in the fall. The profile of 90th percentile concentrations shown in Figure 4-12 is triple-peaked as for some of the other areas, with the highest values in June and July, followed by January, August, and October, and the lowest values in March and September.

Figure 4-11. Distribution of 1999–2002 Days by Season and Severity: Washington

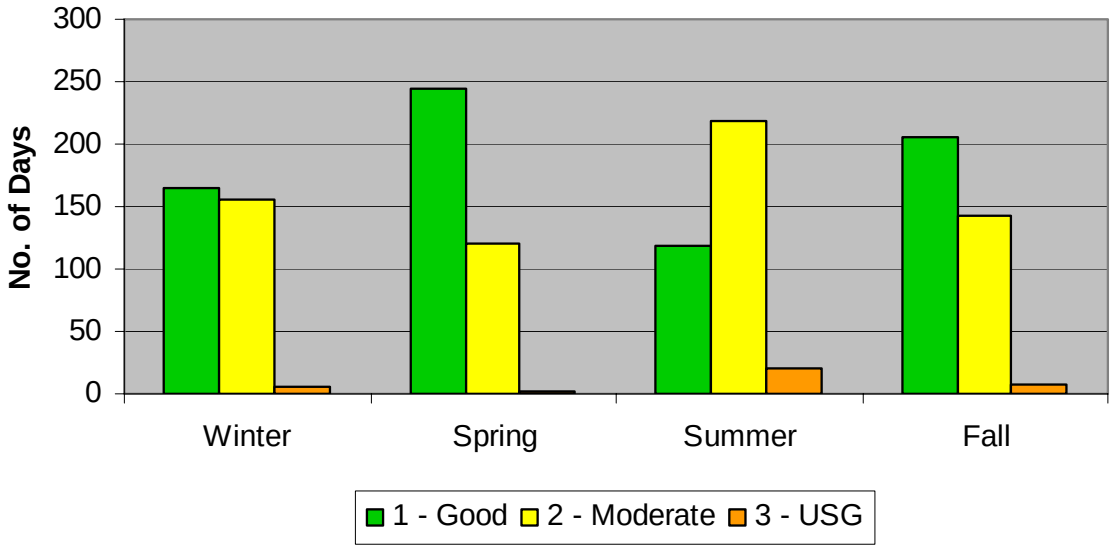
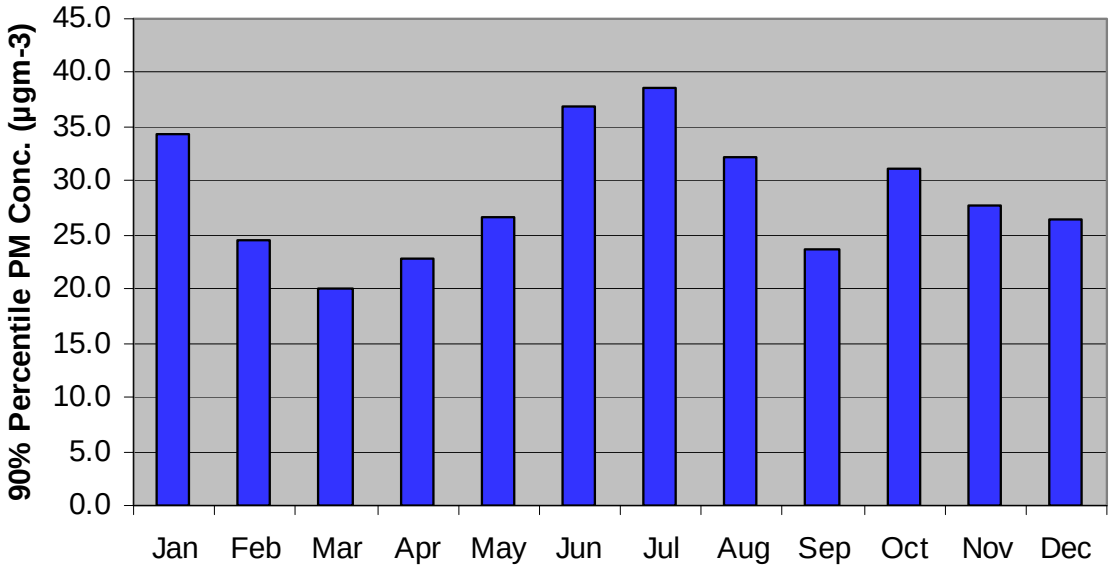


Figure 4-12. 90th Percentile Concentrations by Month (1999–2002): Washington



4.7.2. Meteorological Factors Influencing $PM_{2.5}$ Concentrations

The meteorological conditions associated with the different ranges of $PM_{2.5}$ concentration and specifically the highest PM days for the Washington, D.C. area are discussed in this subsection.

Wind Patterns Associated with High $PM_{2.5}$

Plots comparing the frequency of wind directions and speeds for all days and for low, moderate, and high $PM_{2.5}$ days for the Washington area are presented in Appendix A. The wind information in these plots is for the Dulles Airport (Sterling, VA) upper-air monitoring site. The plots use the same format and contain the same information as for the other areas (described earlier in this section).

The wind roses for Washington (Figures A-9 and A-10) are based on the Dulles Airport sounding data. The upper-level winds are predominately southwesterly to northwesterly for the low PM days, at the time of the morning sounding and west-southwesterly to northerly at the time of the evening sounding. For both sounding times, wind back slightly for the moderate PM days, with a shift to dominant southwesterly winds in the morning and westerly winds in the evening. For the highest PM days, wind speeds are much lower than for the other PM concentration levels and the wind directions are southwesterly, westerly, and northwesterly on the various days.

Categorical Summaries

A comparison of the meteorological characteristics for different ranges of $PM_{2.5}$ concentration in Table 4-14 provides a basis for further distinguishing days within the different categories based on the values of meteorological parameters. In preparing this table, we used the comprehensive meteorological and PM dataset compiled for the CART application. Key meteorological parameters, as used by CART to construct the classification tree, are shaded in this table so that we can focus on the differences in these key parameters as well the differences found

throughout the dataset. Categories 1 to 3 represent the standard three ranges of 24-hour $PM_{2.5}$ concentration: <15.5, 15.5–40.5, and ≥40.5 μgm^{-3}.

Table 4-14. Summary of Mean Air Quality and Meteorological Parameters for Each CART Classification Category: Washington, D.C.

	Category 1	Category 2	Category 3
$PM_{2.5}$ Parameters			
24-hour $PM_{2.5}$ for Washington (μgm^{-3})	10.5	23.0	48.1
Two-days-ago maximum 24-hour $PM_{2.5}$ for Washington, Baltimore, and Gettysburg (μgm^{-3})	16.1	19.3	26.5
Two-days-ago 24-hour $PM_{2.5}$ for Richmond (μgm^{-3})	13.0	15.7	20.9
Surface Meteorological Parameters			
Maximum surface temperature (°C)	17.1	21.8	26.2
Minimum surface temperature (°C)	8.5	12.4	16.8
Surface relative humidity (%)	61.4	68.8	67.6
Surface wind speed (ms^{-1})	3.7	2.6	2.1
Surface wind direction (degrees)	308	235	249
Number of six hour periods with precipitation (range is 1 to 4)	0.3	0.2	0.1
Upper-Air Meteorological Parameters (Dulles Airport)			
850 mb temperature (AM) (°C)	3.7	9.1	13.9
850 mb temperature (PM) (°C)	4.3	9.9	15.4
Temperature gradient (850 mb to surface; AM) (°C)	-3.3	-0.7	-1.8
Temperature gradient (900 mb to surface; AM) (°C)	-1.1	2.2	2.4
	-0.3	3.0	3.9
24-hour difference in 700 mb geopotential height (m)	-3.0	1.0	0.1
Yesterday's 700 mb wind speed (PM) (ms^{-1})	15.2	11.9	7.8
Yesterday's 850 mb wind speed (PM) (ms^{-1})	11.5	8.4	6.1
850 mb wind speed (AM) (ms^{-1})	11.7	8.9	5.9
850 mb wind speed (PM) (ms^{-1})	10.6	9.2	7.4
Yesterday's 700 mb wind direction (PM) (degrees)	285	282	297
Yesterday's 850 mb wind direction (PM) (degrees)	281	276	273
850 mb wind direction (AM) (degrees)	301	277	283
850 mb wind direction (PM) (degrees)	290	266	288
Estimated cloud cover (range of 1 to 3)	1.8	1.9	1.6
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	2	2	2

Table 4-14 summarizes how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for Charlotte. A column-by-column comparison of the values reveals some possible relationships between $PM_{2.5}$ and several of the air quality and meteorological parameters.

High $PM_{2.5}$ in the Washington area is associated with relatively high $PM_{2.5}$ two-days prior—both in the Washington area and in Richmond. Thus, a regional day-to-day build up of $PM_{2.5}$ is indicated for high $PM_{2.5}$ days. Note, however, that neither of the prior-day PM parameters are of high importance to CART.

The surface meteorological parameters indicate a correlation between higher $PM_{2.5}$ concentrations and higher temperatures (primarily reflecting seasonal differences), lower surface wind speeds, and less precipitation. Surface wind directions tend toward southwesterly for the higher ranges of $PM_{2.5}$, compared to northwesterly for the lowest range. Relative humidity is, on average, slightly higher with higher PM.

The upper-air meteorological parameters (based here on the Dulles Airport sounding) indicate that higher $PM_{2.5}$ concentrations occur with higher 850 mb temperatures. There is also a tendency for more stable (positive) lapse rates to be associated with higher $PM_{2.5}$ days. This is especially true for the 900 and 950 mb temperature differences. The difference in geopotential height does not vary regularly across the categories.

Considering the upper-air wind parameters, lower wind speeds aloft characterize the higher PM days. There is no well defined tendency with regard to wind direction aloft, and, on average, westerly winds prevail.

Finally, the cloud cover is less for higher PM, but there and the season parameter does not vary across the three categories.

The input parameters that are most used by CART in the construction of the classification tree (either to define the splits/branching structure) or as surrogates to the primary variables in this regard are highlighted in the table. These include surface temperature, relative humidity, 850 mb temperature, relative humidity, and 950 to surface temperature difference. All of these are also well correlated (directionally) with the $PM_{2.5}$ concentration for the analysis day.

4.7.3. Characteristics of High $PM_{2.5}$ Events

The categorical summary table provides a general overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for the Washington, D.C. area. Within the high $PM_{2.5}$ categories, there are other key differences among the parameters that result in different types of high $PM_{2.5}$ events. We have used the CART results to examine these differences.

Only certain of the CART bins are frequently associated with $PM_{2.5}$ concentrations that are in the USG range or Category 3. Of these, we identified the bins with the most number of days as key bins. Table 4-15 considers the input parameter values for the key USG bins. For Washington, four bins contain 66 percent of the USG days.

Table 4-15. Summary of Mean Air Quality and Meteorological Parameters for Key USG CART Classification Bins: Washington, D.C.

	Bin 34	Bin 11	Bin 19	Bin 37
Number of days	12	4	4	3
$PM_{2.5}$ Parameters				
24-hour $PM_{2.5}$ for Washington (μgm^{-3})	48.1	48.3	50.7	47.4
Two-days-agomaximum 24-hour $PM_{2.5}$ for Washington, Baltimore, and Gettysburg (μgm^{-3})	28.0	14.9	29.7	38.1
Two-days-ago 24-hour $PM_{2.5}$ for Richmond (μgm^{-3})	21.9	13.6	21.6	20.4
Surface Meteorological Parameters				
Maximum surface temperature (°C)	33.3	5.0	27.5	33.7
Minimum surface temperature (°C)	23.2	-1.8	18.5	23.7
Surface relative humidity (%)	61.1	77.2	75.6	66.8
Surface wind speed (ms^{-1})	2.7	1.2	2.1	3.3
Surface wind direction (degrees)	259	45	225	225
Number of six hour periods with precipitation (range is 1 to 4)	0.0	0.3	0.0	0.3
Upper-Air Meteorological Parameters (Dulles Airport)				
850 mb temperature (AM) (°C)	17.2	-1.0	14.7	17.3
850 mb temperature (PM) (°C)	19.1	2.9	14.7	18.8
Temperature gradient (850 mb to surface; AM) (°C)	-2.5	3.4	-3.7	-5.4
Temperature gradient (900 mb to surface; AM) (°C)	2.1	3.7	2.2	-1.4
Temperature gradient (950 mb to surface; AM) (°C)	4.0	3.2	3.7	0.3
24-hour difference in 700 mb geopotential height (m)	-0.5	-5.8	-3.2	-26.8
Yesterday's 700 mb wind speed (PM) (ms^{-1})	7.9	10.9	3.5	8.1
Yesterday's 850 mb wind speed (PM) (ms^{-1})	6.0	6.4	5.5	5.5
850 mb wind speed (AM) (ms^{-1})	5.1	7.0	5.3	12.1
850 mb wind speed (PM) (ms^{-1})	6.7	10.8	8.0	6.5
Yesterday's 700 mb wind direction (PM) (degrees)	315	270	315	243
Yesterday's 850 mb wind direction (PM) (degrees)	301	288	225	207
850 mb wind direction (AM) (degrees)	306	243	243	315
850 mb wind direction (PM) (degrees)	286	270	270	333
Estimated cloud cover (range of 1 to 3)	1.5	1.5	2.0	2.0
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	3	1	2	3

While many of the characteristics are similar for the exceedance bins, there are some differences. These provide insight into the factors influencing the high PM days within each bin.

Days within Bins 34 and 19 have similar values for previous day PM concentration, whereas Bin 11 days have much lower values, on average, and Bin 37 days have much higher values, on average, in the local Washington area than days within the other bins. Thus these bins are characterized by regional-scale build up of PM (Bins 34 and 19), rapid build up of PM (Bin 11), and persistent high values in the local area (Bin 37). From the temperatures, as well as from the seasonal index, the bins represent different times of the year—with Bin 11 for winter days, Bin 19 for transitional season days, and Bins 34 and 37 for summer days.

In addition to the lowest temperatures and prior-day PM values, the days within Bin 11 are characterized by the lowest surface wind speeds and the deepest stable layers. Surface wind directions from the northeast are also unique to this bin.

Bin 19 is comprised primarily of transitional season days and these days have the second lowest wind speeds, on average, but otherwise conditions that tend to be intermediate to the other bins.

The two bins comprised mostly of summer days have slightly higher surface wind speeds, on average, and lower relative humidity, than the other two bins. They differ from one another in the stability characteristics such that days within Bin 37 are much less stable. Days within this bin also show decreasing heights and high wind speeds aloft during the morning hours, compared to days within Bin 34. Thus, there appear to be two different summertime regimes with different synoptic characteristics.

Next, we examine the conditions associated with each day or episode.

Data retrieval and availability for the Washington area were high for the period 1999–2002, and thirty five orange days occurred during this period. Of all the areas of interest in the MARAMA region, the Washington area experienced the largest number of USG days during this period. The specific dates, including the observed $PM_{2.5}$ concentration (μgm^{-3}), are listed in Table 4-16.

Table 4-16. USG days for Washington, D.C.: 1999–2002

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG Day for Other Areas?
July 22, 1999	Thursday	37	56.3	
September 27, 1999	Monday	19	67.0	
January 1, 2000	Saturday	26	46.0	Charlotte
June 10, 2000	Saturday	34	42.1	Baltimore, Philadelphia, Newark
June 23, 2000	Friday	26	47.7	
July 1, 2000	Saturday	23	47.0	
July 9, 2000	Sunday	34	41.2	
October 16, 2000	Monday	12	40.5	
October 26, 2000	Thursday	17	50.3	Baltimore
October 27, 2000	Friday	19	44.3	
November 8, 2000	Wednesday	12	47.0	Baltimore
November 9, 2000	Thursday	22	42.1	
January 8, 2001	Monday	11	48.2	
January 13, 2001	Saturday	11	49.4	Baltimore

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG Day for Other Areas?
January 18, 2001	Thursday	11	40.6	
January 23, 2001	Tuesday	11	54.9	Baltimore, Wilmington, Philadelphia
January 24, 2001	Wednesday	10	49.4	Baltimore, Wilmington, Philadelphia
May 4, 2001	Friday	19	41.0	Baltimore
June 13, 2001	Wednesday	26	46.1	Baltimore
June 29, 2001	Friday	19	50.5	Baltimore, Wilmington, Philadelphia
August 6, 2001	Monday	37	43.8	Baltimore, Wilmington, Philadelphia, Newark
August 7, 2001	Tuesday	34	44.8	
August 8, 2001	Wednesday	34	48.7	Roanoke, Baltimore, Richmond, Wilmington
August 9, 2001	Thursday	34	50.5	Baltimore, Richmond, Roanoke, Wilmington, Philadelphia
November 16, 2001	Friday	12	45.2	
June 25, 2002	Tuesday	34	56.1	Baltimore, Wilmington
July 2, 2002	Tuesday	34	55.5	Baltimore, Richmond, Wilmington, Philadelphia, Newark
July 3, 2002	Wednesday	34	49.7	Baltimore, Richmond, Wilmington, Philadelphia
July 4, 2002	Thursday	30	59.1	
July 7, 2002	Sunday	34	43.6	
July 8, 2002	Monday	34	49.5	
July 9, 2002	Tuesday	37	42.0	Philadelphia
July 18, 2002	Thursday	34	52.2	Baltimore, Richmond, Charlotte, Wilmington, Newark
July 19, 2002	Friday	34	43.1	Baltimore
August 13, 2002	Tuesday	31	47.8	Baltimore, Richmond, Wilmington, Philadelphia

For the Washington area, the greatest number of USG days occurs in the summer months, but overall the days are distributed among all quarters of the year. The observed concentration levels on the USG days fall into the low to mid-range for the category, with the exception of the 67 μgm^{-3} measured on September 27, 1999, which actually falls into the red category. Given the number of orange days for the Washington area, rather than discuss each orange day individually, the discussion will include groups of days by season, or specific multi-day episodes.

For the USG days measured during the summer months, the Washington area experiences similar meteorological conditions that lead to high PM concentrations: light winds, high temperatures, limited mixing, high humidity, and high solar radiation. Important features that influence regional PM formation are the location and strength of the upper level ridges that affect the regional wind, temperature, stability fields, as well as the cloud and precipitation fields, which are important influences on solar radiation and its role in the photochemistry of

PM formation. Another important aspect is the location and strength of the surface high-pressure system and the resulting influence on surface winds, temperatures, cloud cover, humidity, precipitation, and local dispersion characteristics. On many of the observed summer USG days for the Washington area, the upper-level ridge is located directly over the area or is very weak, reflecting typical summer conditions in the upper atmosphere. With an upper-level ridge in this position, the temperatures aloft increase and the wind speeds decrease, leading to a buildup of PM over multiple days. On many of these days, skies are relatively clear (hazy) and precipitation is also suppressed in the area, which allows for further buildup of PM. The August 6–9, 2001 and the July 1–4, 2002 periods exhibited multiple USG days throughout the region and are good examples of widespread, persistent summertime conditions that lead to high PM in the area.

As noted above, observed USG days for the Washington area occur in every quarter of the year. The summer months experience the highest PM concentrations in the MARAMA region (a large portion of this being sulfate) because of the enhancement in sulfate formation due to photochemistry and the availability of moisture compared to drier wintertime conditions.

The wintertime conditions for observed USG days in the Washington area, such as those that occurred during January 2000 and 2001 indicate a number of features are important in influencing the buildup of PM concentrations. The locations of the upper-level ridges (and troughs) that migrate across the area in the winter months influence the strength of the surface features. During January 2001, for example, the Washington area (and the entire East Coast) was under the influence of a cold air mass from Canada. This air mass was associated with a subsidence aloft, and a strong surface high-pressure system, which resulted in inversions throughout the area that limited dispersion and allowed PM concentrations to build up over the area. These conditions persisted until the upper level features moved across the area, bringing unsettled weather, precipitation, and other conditions not conducive to a build up of PM. The January 23–24, 2001 period is a good example of widespread, persistent wintertime conditions leading to high observed PM at multiple sites throughout the region.

During the spring and fall months of the year, regional weather patterns that limit wind speeds and dispersion occur in the Washington area and, on occasion, are enough to result in high PM concentrations that fall into the USG category.

This review of the meteorological conditions indicates the high PM concentrations occur under a variety of synoptic situations, and that these vary by season. Interestingly, the CART-based classification strongly replicates this and most days within the key high PM bins correspond to the same seasonal periods.

The results for this area are a good example of how very different conditions can lead to high PM concentrations. In this case, the categorical summaries should not be used to guide the forecasting, and instead the bin by bin characteristics must be considered.

4.8. Factors Influencing $PM_{2.5}$ Concentrations for Baltimore, MD

The area-wide maximum $PM_{2.5}$ for the Baltimore area was defined for this study as the maximum value over all of the sites listed as the local Baltimore sites in Table 2-1.

4.8.1. Summary of Observed $PM_{2.5}$ Data (1999–2002)

The Baltimore-area daily maximum PM2.5 variable was defined as the maximum over fourteen FRM sites in Anne Arundel and Harford Counties as well as the city of Baltimore. Data from two additional FRM monitors were also used whenever data were missing from their collocated monitors in the primary set of fourteen. Figure 4-13 shows how days of different PM severity are distributed over the seasons. Although USG days appear in all seasons, they most often occur in the summer, when most days are moderate or worse. Overall, three percent of the days are USG and about half of these occur in summertime, and another quarter in winter. Figure 4-14 shows the fine mass concentrations at the 90th percentile for each month. The summer months are high, as one would expect, but the highest 90th percentile value actually occurs in January.

Figure 4-13. Distribution of 1999–2002 Days by Season and Severity: Baltimore

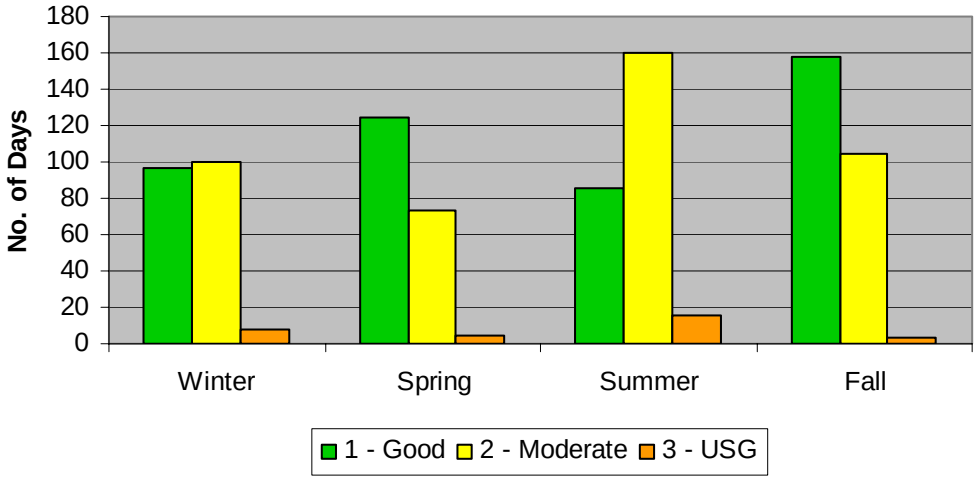
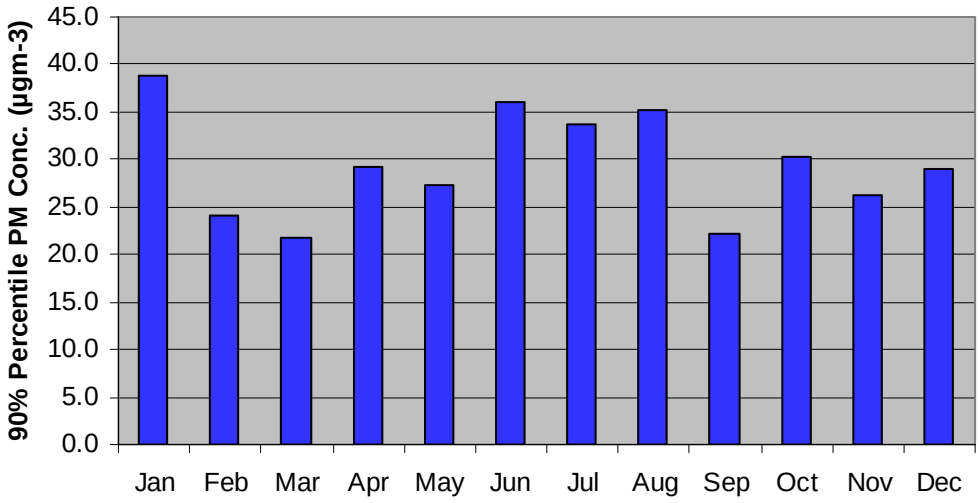


Figure 4-14. 90th Percentile Concentrations by Month (1999–2002): Baltimore



4.8.2. Meteorological Factors Influencing $PM_{2.5}$ Concentrations

The meteorological conditions associated with the different ranges of $PM_{2.5}$ concentration and specifically the highest PM days for the Baltimore area are discussed in this subsection.

Wind Patterns Associated with High $PM_{2.5}$

Plots comparing the frequency of wind directions and speeds for all days and for low, moderate, and high $PM_{2.5}$ days for the Baltimore area are presented in Appendix A. The wind information in these plots is for the Dulles Airport (Sterling, VA) upper-air monitoring site. The plots use the same format and contain the same information as for the other areas (described earlier in this section).

The wind roses for Baltimore (Figures A-11 and A-12) are based on the Dulles Airport sounding data. The upper-level winds are predominately westerly to northwesterly for the low PM days, at the time of both the morning and evening soundings, but there are some days with southwesterly winds during the evening hours in this category. For both sounding times, wind directions, on average, back to a more southwesterly direction for the moderate PM days, with lower wind speeds than for the lower PM days. For the highest PM days, wind speeds are much lower than for the other PM concentration levels and the wind directions are west-southwesterly to northwesterly at the time of the morning sounding and southerly to northwesterly at the time of the evening sounding.

Categorical Summaries

A comparison of the meteorological characteristics for different ranges of $PM_{2.5}$ concentration in Table 4-17 provides a basis for further distinguishing days within the different categories based on the values of meteorological parameters. In preparing this table, we used the comprehensive meteorological and PM dataset compiled for the CART application. Key meteorological parameters, as used by CART to construct the classification tree, are shaded in this table so that we can focus on the differences in these key parameters as well the differences found throughout the dataset. Categories 1 to 3 represent the standard three ranges of 24-hour $PM_{2.5}$ concentration: <15.5, 15.5–40.5, and ≥40.5 μgm^{-3}.

Table 4-17. Summary of Mean Air Quality and Meteorological Parameters for Each CART Classification Category: Baltimore

	Category 1	Category 2	Category 3
$PM_{2.5}$ Parameters			
24-hour $PM_{2.5}$ for Baltimore (μgm^{-3})	10.5	23.2	49.2
Two-days-ago maximum 24-hour $PM_{2.5}$ for Washington, Baltimore, and Gettysburg (μgm^{-3})	17.1	18.9	26.5
Two-days-ago 24-hour $PM_{2.5}$ for Richmond (μgm^{-3})	13.6	15.5	19.4
Surface Meteorological Parameters			
Maximum surface temperature (°C)	16.8	21.7	24.1
Minimum surface temperature (°C)	7.0	10.3	12.9
Surface relative humidity (%)	64.1	70.4	69.9
Surface wind speed (ms^{-1})	3.0	1.9	1.7
Surface wind direction (degrees)	278	218	202
Number of six hour periods with precipitation (range is 1 to 4)	0.3	0.2	0.1
Upper-Air Meteorological Parameters (Dulles Airport)			
850 mb temperature (AM) (°C)	4.1	9.3	13.4
850 mb temperature (PM) (°C)	4.7	10.2	14.3
Temperature gradient (850 mb to surface; AM) (°C)	-3.5	-0.7	-0.3
Temperature gradient (900 mb to surface; AM) (°C)	-1.2	2.2	3.1
Temperature gradient (950 mb to surface; AM) (°C)	-0.3	3.1	3.8
24-hour difference in 700 mb geopotential height (m)	-5.1	1.9	-1.8
Yesterday's 700 mb wind speed (PM) (ms^{-1})	15.4	11.9	8.5
Yesterday's 850 mb wind speed (PM) (ms^{-1})	11.4	8.3	6.5
850 mb wind speed (AM) (ms^{-1})	11.1	8.6	6.6
850 mb wind speed (PM) (ms^{-1})	10.5	9.0	7.8
Yesterday's 700 mb wind direction (PM) (degrees)	282	282	295
Yesterday's 850 mb wind direction (PM) (degrees)	279	276	267
850 mb wind direction (AM) (degrees)	296	275	289
850 mb wind direction (PM) (degrees)	292	266	262
Estimated cloud cover (range of 1 to 3)	1.9	1.9	1.7
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	2	2	2

Table 4-17 provides a general overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for Baltimore. A column-by-column comparison of the values reveals some clear tendencies in several of the air quality and meteorological parameters.

High $PM_{2.5}$ in the Baltimore area is clearly associated with relatively high $PM_{2.5}$ two-days prior—in the Baltimore-Washington area and to a lesser extent the Richmond area. Thus, a regional day-to-day build up of $PM_{2.5}$ is indicated for high $PM_{2.5}$ days.

The surface meteorological parameters indicate a correlation between higher $PM_{2.5}$ concentrations and higher temperatures (primarily reflecting seasonal differences), lower surface wind speeds, and less precipitation. Surface wind directions tend toward southerly (from westerly) for the higher ranges of $PM_{2.5}$. Relative humidity is slightly higher, on average, for the higher PM categories.

The upper-air meteorological parameters (based here on the Dulles Airport sounding) indicate that higher $PM_{2.5}$ concentrations occur with higher 850 mb temperatures. There is also a tendency for more stable (positive) lapse rates to be associated with higher $PM_{2.5}$ days. This is especially true for the 900 and 950 mb temperature differences.

Lower wind speeds aloft also distinguish the higher $PM_{2.5}$ concentration days. There is no pronounced difference in average wind direction among the categories.

Finally, the cloud cover and season parameters do not vary much across the three categories.

The input parameters that are most used by CART in the construction of the classification tree (either to define the splits/branching structure) or as surrogates to the primary variables in this regard are highlighted in the table. These include two-days-ago $PM_{2.5}$ for the Baltimore-Washington area, surface temperature, surface wind speed, 850 mb temperature, the 900 to surface temperature difference, and 850 mb wind speed at the time of the morning sounding. All of these vary regularly with $PM_{2.5}$ concentration for the analysis day.

4.8.3. Characteristics of High $PM_{2.5}$ Events

The categorical summary table provides a general overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for the Baltimore area. Within the high $PM_{2.5}$ categories, there are other key differences among the parameters that result in different types of high $PM_{2.5}$ events. We have used the CART results to examine these differences.

Only certain of the CART bins are frequently associated with $PM_{2.5}$ concentrations that are in the USG range or Category 3. Of these, we identified the bins with the most number of days as key bins. Table 4-18 considers the input parameter values for the key USG bins. For Baltimore there are four key bins and these contain 77 percent of the USG days.

Table 4-18. Summary of Mean Air Quality and Meteorological Parameters for Key USG CART Classification Bins: Baltimore

	Bin 29	Bin 16	Bin 34	Bin 18
Number of days	13	4	4	3
$PM_{2.5}$ Parameters				
24-hour $PM_{2.5}$ for Baltimore (μgm^{-3})	49.0	49.6	53.3	51.0
Two-days-ago maximum 24-hour $PM_{2.5}$ for Washington, Baltimore, and Gettysburg (μgm^{-3})	32.7	22.7	19.6	18.9
Two-days-ago 24-hour $PM_{2.5}$ for Richmond (μgm^{-3})	21.8	17.5	15.3	16.4

	Bin 29	Bin 16	Bin 34	Bin 18
Surface Meteorological Parameters				
Maximum surface temperature (°C)	34.1	5.0	32.9	17.2
Minimum surface temperature (°C)	21.2	-7.6	21.9	9.6
Surface relative humidity (%)	63.1	76.3	73.6	83.9
Surface wind speed (ms^{-1})	1.8	0.4	1.7	2.1
Surface wind direction (degrees)	230	45	225	90
Number of six hour periods with precipitation (range is 1 to 4)	0.0	0.0	0.3	0.0
Upper-Air Meteorological Parameters (Dulles Airport)				
850 mb temperature (AM) (°C)	18.5	1.9	17.8	11.9
850 mb temperature (PM) (°C)	18.2	3.3	18.1	12.9
Temperature gradient (850 mb to surface; AM) (°C)	-1.6	11.3	-2.7	-0.7
Temperature gradient (900 mb to surface; AM) (°C)	3.6	11.8	0.3	1.5
Temperature gradient (950 mb to surface; AM) (°C)	5.4	8.0	2.0	2.9
24-hour difference in 700 mb geopotential height (m)	3.1	-12.3	-24.5	16.2
Yesterday's 700 mb wind speed (PM) (ms^{-1})	5.3	7.3	13.9	12.3
Yesterday's 850 mb wind speed (PM) (ms^{-1})	4.9	6.2	6.0	10.6
850 mb wind speed (AM) (ms^{-1})	4.5	9.1	6.0	7.4
850 mb wind speed (PM) (ms^{-1})	5.1	7.8	11.7	9.6
Yesterday's 700 mb wind direction (PM) (degrees)	321	270	297	270
Yesterday's 850 mb wind direction (PM) (degrees)	270	288	270	243
850 mb wind direction (AM) (degrees)	306	270	270	297
850 mb wind direction (PM) (degrees)	279	252	270	243
Estimated cloud cover (range of 1 to 3)	1.7	1.3	1.8	2.0
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	3	1	3	2

While many of the characteristics are similar for the high PM bins, there are some differences. These provide insight into the factors influencing the concentration that characterize each bin.

Days within Bin 29 are characterized by the highest two-days-ago day PM concentrations, whereas days within the other three key bins have lower and more consistent values, both for the Baltimore-Washington and Richmond areas. From the temperatures, as well as from the seasonal index, the bins represent different times of the year—with Bin 16 for winter days, Bin 18 for transitional season days, and Bins 29 and 34 for summer days.

In addition to the lowest temperatures, the days within Bin 16 are characterized by very stable temperature differences that are much larger than for the other key bins. The stable layer is also

deeper for this bin and extends through the 850 mb level. Days within this bin have the lowest wind speeds overall, with an average that is nearly zero. Surface wind directions from the northeast are also unique to this bin.

Bin 18 is comprised primarily of transitional season days. Wind speeds tend to be higher, on average, than for the other bins, both near the surface and aloft. The change in geopotential height is most positive for days within this bin. Surface wind directions are, on average, from the east, which is unique to this bin. Cloud cover is the greatest over all key bins.

The two bins comprised mostly of summer days have higher temperatures and intermediate surface wind speeds when compared to the other key bins. Days within these bins also exhibit southwesterly surface wind directions. Relative humidity is higher for Bin 34. The bins also differ from one another in the stability characteristics such that days within Bin 34 are less stable. Days within this bin also show decreasing heights and high wind speeds aloft during the morning hours, compared to days within Bin 29. Thus, there appear to be two different summertime regimes with different synoptic characteristics.

Next, we examine the conditions associated with each day or episode.

Data retrieval and availability for the Baltimore area were relatively high (although not as high as for Washington) for the period 1999–2002, and 31 USG days occurred during this period, resulting in the second largest number of orange days during this period in the MARAMA region. The specific dates, including the observed $PM_{2.5}$ concentration ($\mu g/m^3$), are as follows:

Table 4-19. USG Days for Baltimore: 1999–2002

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG Day for Other Areas?
June 2, 2000	Friday	34	43.3	Bristol
June 11, 2000	Sunday	29	42.0	Washington, Bristol
October 26, 2000	Thursday	18	53.4	Washington
November 8, 2000	Wednesday	13	41.3	Washington
December 16, 2000	Saturday	15	50.4	
January 12, 2001	Friday	3	41.4	
January 13, 2001	Saturday	16	53.6	Washington
January 14, 2001	Sunday	16	45.5	
January 23, 2001	Tuesday	16	56.7	Washington, Wilmington, Philadelphia
January 24, 2001	Wednesday	5	43.2	Washington, Wilmington, Philadelphia
January 25, 2001	Thursday	12	63.7	
April 7, 2001	Saturday	18	52.5	
April 10, 2001	Tuesday	18	47.1	
May 3, 2001	Thursday	29	40.9	
May 4, 2001	Friday	29	43.8	Washington
June 12, 2001	Tuesday	28	40.6	Washington
June 28, 2001	Thursday	29	42.9	Wilmington, Philadelphia
June 29, 2001	Friday	29	62.1	Washington, Wilmington, Philadelphia

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG Day for Other Areas?
August 5, 2001	Sunday	29	51.3	
August 6, 2001	Monday	29	45.1	Washington, Wilmington, Newark
August 8, 2001	Wednesday	29	46.7	Washington, Richmond, Roanoke, Wilmington
August 9, 2001	Thursday	29	53.4	Washington, Richmond, Wilmington, Philadelphia
June 25, 2002	Tuesday	34	59.6	Washington
July 2, 2002	Tuesday	29	54.1	Washington, Richmond, Wilmington, Philadelphia, Newark
July 3, 2002	Wednesday	29	50.7	Washington, Richmond, Wilmington, Philadelphia
July 18, 2002	Thursday	29	50.5	Washington, Richmond, Charlotte, Wilmington, Newark
July 19, 2002	Friday	34	46.3	Washington, Richmond
August 13, 2002	Tuesday	29	52.7	Washington, Richmond
August 24, 2002	Saturday	34	64.2	
October 4, 2002	Friday	26	41.8	
December 10, 2002	Tuesday	16	42.7	

The high PM days in the Baltimore area during this period are distributed more evenly across the seasons than for Washington. Although high PM occurs more often in the summer months, high PM days occurred during all quarters of the year. Due to their proximity, the Baltimore and Washington areas encounter very similar weather conditions leading to high PM concentrations throughout the year. The January 23–24, 2001 wintertime conditions leading to high PM in Washington also caused high PM in the Baltimore area, extending to Wilmington and Philadelphia as well.

A rather severe summertime episode occurred during the period August 5–9, 2001. This episode was dominated by a large upper-level ridge extending over the entire U.S. with a strong surface high-pressure system centered over the mid-Atlantic states. This pattern persisted for several days. High temperatures were in the upper 80's at the beginning of the period to near 100 at the end of the period. Winds at the upper levels were light and westerly, while surface winds were light and variable. Skies were reported hazy during this period with partly cloudy conditions and little precipitation. The combination of persistent stagnant conditions led to a regional buildup of PM throughout the MARAMA region with USG days reported at seven of the nine areas of interest during one or more days of this episode.

The results for this area are another example of how different conditions can lead to high PM concentrations. CART effective separates into different types of high PM events that share seasonal characteristics and then separates them further into bins based on other differences in the parameters. The is not just one pathway to high $PM_{2.5}$. In this case, the categorical summaries should not be used to guide the forecasting, and instead the bin by bin characteristics must be considered.

4.9. Factors Influencing $PM_{2.5}$ Concentrations for Philadelphia, PA

The area-wide maximum $PM_{2.5}$ for the Philadelphia area was defined for this study as the maximum value over all of the sites listed as the local Philadelphia sites in Table 2-1.

4.9.1. Summary of Observed $PM_{2.5}$ Data (1999–2002)

Five FRM monitors in the greater Philadelphia area determine the area-wide maximum $PM_{2.5}$ concentrations. Two percent of the days with available data were USG, and as Figure 4-15 shows, most of these days, as usual, appeared in the summer, although six days, or one-third the summer total, appeared in winter. Unlike the other areas, good and moderate summer days are closely matched in quantity, and good days are in the minority in winter. Figure 4-16 shows the 90th percentile concentrations, which are highest in June but about equal in January and July, which share the second-highest rank.

Figure 4-15. Distribution of 1999–2002 Days by Season and Severity: Philadelphia

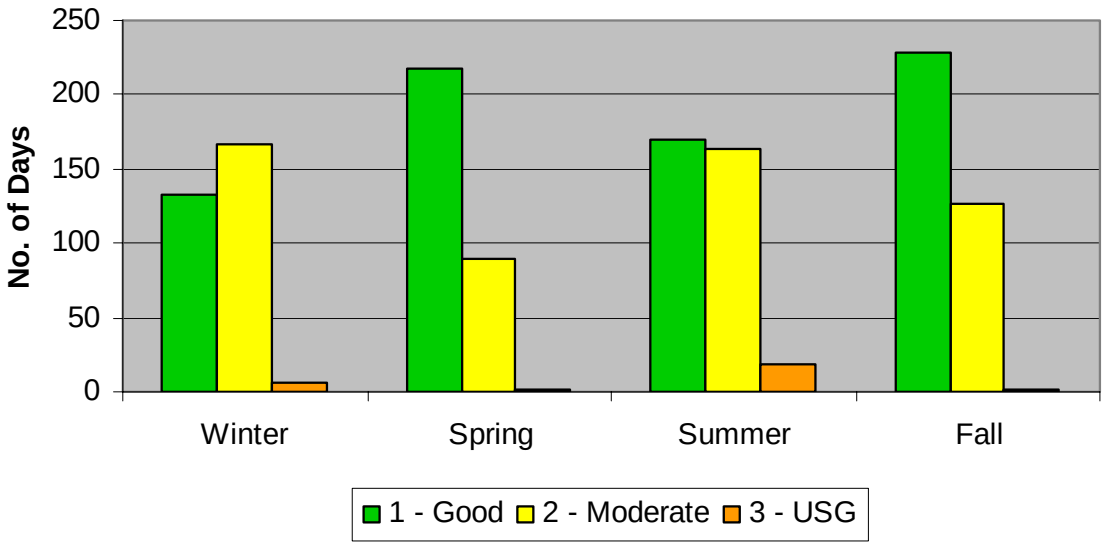
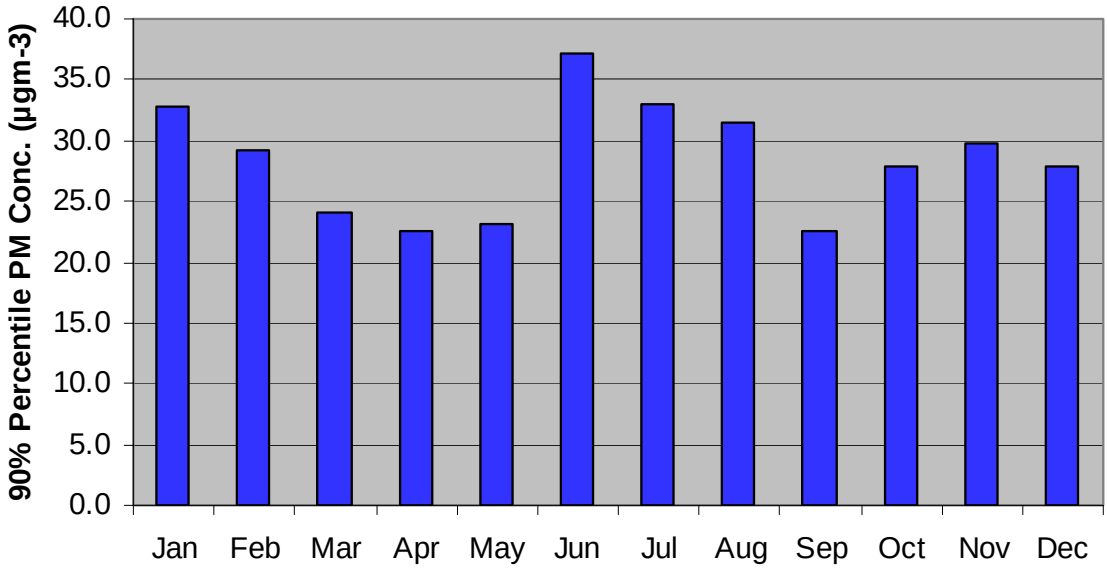


Figure 4-16. 90th Percentile Concentrations by Month (1999–2002): Philadelphia



4.9.2. Meteorological Factors Influencing $PM_{2.5}$ Concentrations

The meteorological conditions associated with the different ranges of $PM_{2.5}$ concentration and specifically the highest PM days for the Philadelphia area are discussed in this subsection.

Wind Patterns Associated with High $PM_{2.5}$

Plots comparing the frequency of wind directions and speeds for all days and for low, moderate, and high $PM_{2.5}$ days for the Philadelphia area are presented in Appendix A. The wind information in these plots is for the Dulles Airport (Sterling, VA) upper-air monitoring site. The plots use the same format and contain the same information as for the other areas (described earlier in this section).

The wind roses for Philadelphia (Figures A-13 and A-14) are based on the Dulles Airport sounding data. The upper-level winds are predominately westerly to northwesterly for the low PM days, at the time of both the morning and evening soundings. Northwesterly winds characterize the greatest number of days. For both sounding times, wind directions, on average, back to a more southwesterly direction for the moderate PM days, with lower wind speeds than for the lower PM days. For the highest PM days, wind speeds are much lower than for the other PM concentration levels and the wind directions range from southwesterly to northwesterly; wind predominantly westerly wind directions at the time of the morning sounding and predominantly southwesterly wind directions at the time of the evening sounding.

Categorical Summaries

A comparison of the meteorological characteristics for different ranges of $PM_{2.5}$ concentration in Table 4-20 provides a basis for further distinguishing days within the different categories based on the values of meteorological parameters. In preparing this table, we used the comprehensive meteorological and PM dataset compiled for the CART application. Key meteorological parameters, as used by CART to construct the classification tree, are shaded in this table so that we can focus on the differences in these key parameters as well the differences found

throughout the dataset. Categories 1 to 3 represent the standard three ranges of 24-hour $PM_{2.5}$ concentration: <15.5, 15.5–40.5, and ≥40.5 μgm^{-3}.

Table 4-20. Summary of Mean Air Quality and Meteorological Parameters for Each CART Classification Category: Philadelphia

	Category 1	Category 2	Category 3
$PM_{2.5}$ Parameters			
24-hour $PM_{2.5}$ for Philadelphia (μgm^{-3})	9.8	22.9	46.8
Two-days-ago maximum 24-hour $PM_{2.5}$ for Camden and New Castle (μgm^{-3})	15.5	17.5	26.7
Two-days-ago 24-hour $PM_{2.5}$ for Washington, Baltimore, and Gettysburg (μgm^{-3})	16.9	18.6	30.2
Surface Meteorological Parameters			
Maximum surface temperature (°C)	17.2	19.4	26.7
Minimum surface temperature (°C)	8.9	9.9	15.5
Surface relative humidity (%)	62.8	69.7	68.8
Surface wind speed (ms^{-1})	3.9	3.0	2.8
Surface wind direction (degrees)	274	188	191
Number of six hour periods with precipitation (range is 1 to 4)	0.3	0.2	0.1
Upper-Air Meteorological Parameters (Dulles Airport)			
850 mb temperature (AM) (°C)	4.9	7.9	14.1
850 mb temperature (PM) (°C)	5.6	8.7	14.7
Temperature gradient (850 mb to surface; AM) (°C)	-3.4	-0.4	0.3
Temperature gradient (900 mb to surface; AM) (°C)	-1.2	2.6	4.1
Temperature gradient (950 mb to surface; AM) (°C)	-0.3	3.4	5.5
24-hour difference in 700 mb geopotential height (m)	-1.3	-1.5	-18.7
Yesterday's 700 mb wind speed (PM) (ms^{-1})	14.6	12.4	7.7
Yesterday's 850 mb wind speed (PM) (ms^{-1})	11.0	8.8	6.1
850 mb wind speed (AM) (ms^{-1})	10.9	9.7	6.9
850 mb wind speed (PM) (ms^{-1})	10.2	9.8	7.6
Yesterday's 700 mb wind direction (PM) (degrees)	244	242	263
Yesterday's 850 mb wind direction (PM) (degrees)	245	224	227
850 mb wind direction (AM) (degrees)	256	234	232
850 mb wind direction (PM) (degrees)	248	222	215
Estimated cloud cover (range of 1 to 3)	1.9	1.8	1.6
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	2	2	2

Table 4-20 provides an overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for Philadelphia.

High $PM_{2.5}$ in the Philadelphia area is associated with relatively high $PM_{2.5}$ two-days prior—in both the Philadelphia (Camden-Wilmington) and Baltimore-Washington areas. Thus, a regional day-to-day build up of $PM_{2.5}$ is indicated for high $PM_{2.5}$ days.

The surface meteorological parameters indicate a correlation between higher $PM_{2.5}$ concentrations and higher temperatures (primarily reflecting seasonal differences) and less precipitation. Surface wind directions tend toward southerly for the higher ranges of $PM_{2.5}$, compared to westerly for the lowest range of concentration. There is no clear tendency for relative humidity and surface wind speed.

The upper-air meteorological parameters (based here on the Dulles Airport sounding) indicate that higher $PM_{2.5}$ concentrations occur with higher 850 mb temperatures and more stable (positive) lapse rates. The difference in geopotential height is much more negative for the higher PM days.

Considering the upper-air wind data, the higher PM days are characterized by lower wind speeds aloft. Winds aloft are, on average, southwesterly, for all three categories.

High PM is associated with slightly less cloud cover; overall, the season parameters do not distinguish the categories at this most general level.

The input parameters that are most used by CART in the construction of the classification tree (either to define the splits/branching structure) or as surrogates to the primary variables in this regard are highlighted in the table. These include two-days-ago for the Camden-New Castle area, surface temperature, 850 mb temperature, and 900 to surface temperature difference. All of these are also well correlated with the $PM_{2.5}$ concentration for the analysis day.

4.9.3. Characteristics of High $PM_{2.5}$ Events

The categorical summary table provides a general overview of how average conditions vary across (and potentially lead to) different 8-hour ozone concentration levels for the Philadelphia area. Within the high $PM_{2.5}$ categories, there are other key differences among the parameters that result in different types of high $PM_{2.5}$ events. We have used the CART results to examine these differences.

Only certain of the CART bins are frequently associated with $PM_{2.5}$ concentrations that are in the USG range or Category 3. Of these, we identified the bins with the most number of days as key bins. Table 4-21 considers the input parameter values for the key USG bins. For Philadelphia, there are two key bins containing 17 and 7, respectively, of the 28 USG days.

Table 4-21. Summary of Mean Air Quality and Meteorological Parameters for Key USG CART Classification Bins: Philadelphia.

	Bin 34	Bin 26
Number of days	17	7
$PM_{2.5}$ Parameters		
24-hour $PM_{2.5}$ for Philadelphia (μgm^{-3})	46.2	47.3
Two-days-ago maximum 24-hour $PM_{2.5}$ for Camden and New Castle (μgm^{-3})	30.9	20.0

	Bin 34	Bin 26
Two-days-ago 24-hour $PM_{2.5}$ for Washington, Baltimore, and Gettysburg (μgm^{-3})	33.9	26.7
Surface Meteorological Parameters		
Maximum surface temperature (°C)	35.0	12.5
Minimum surface temperature (°C)	23.9	0.8
Surface relative humidity (%)	65.1	77.0
Surface wind speed (ms^{-1})	3.5	1.3
Surface wind direction (degrees)	184	214
Number of six hour periods with precipitation (range is 1 to 4)	0.2	0.0
Upper-Air Meteorological Parameters (Dulles Airport)		
850 mb temperature (AM) (°C)	18.4	7.1
850 mb temperature (PM) (°C)	19.1	7.1
Temperature gradient (850 mb to surface; AM) (°C)	-3.3	8.3
Temperature gradient (900 mb to surface; AM) (°C)	1.0	10.7
Temperature gradient (950 mb to surface; AM) (°C)	3.2	10.2
24-hour difference in 700 mb geopotential height (m)	-14.1	-24.8
Yesterday's 700 mb wind speed (PM) (ms^{-1})	6.7	8.4
Yesterday's 850 mb wind speed (PM) (ms^{-1})	5.6	6.6
850 mb wind speed (AM) (ms^{-1})	5.7	9.9
850 mb wind speed (PM) (ms^{-1})	6.6	10.7
Yesterday's 700 mb wind direction (PM) (degrees)	260	270
Yesterday's 850 mb wind direction (PM) (degrees)	238	217
850 mb wind direction (AM) (degrees)	232	214
850 mb wind direction (PM) (degrees)	221	189
Estimated cloud cover (range of 1 to 3)	1.8	1.1
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	3	1

The two key high PM bins represent winter and summer types of PM events.

Days within Bin 26 (containing a majority of winter time days) are associated with lower two-days-ago PM concentrations, yet higher concentrations, on average, on the analysis days, compared to days within Bin 34 (the summertime bin). Temperatures and surface wind speeds are much lower for days within Bin 26. The days within this bin are also distinguished by very stable lapse rates and a deep stable layer. Wind speeds aloft are greater for Bin 26 than for Bin 34, as expected during wintertime synoptic conditions.

Data retrieval and availability for the Philadelphia area were high for the period 1999–2002, and 28 USG days occurred during this period. The specific dates, including the observed $PM_{2.5}$ concentration (μgm^{-3}), are listed in Table 4-22.

Table 4-22. USG Days for Philadelphia: 1999–2002

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG Day for Other Areas?
July 19, 1999	Monday	34	50.5	Wilmington
July 24, 1999	Saturday	34	46.9	Wilmington
July 31, 1999	Saturday	34	42.3	Wilmington
October 30, 1999	Saturday	26	41.8	
February 4, 2000	Friday	2	49.2	
February 10, 2000	Thursday	26	48.9	Wilmington
February 11, 2000	Friday	26	48.0	
March 9, 2000	Thursday	26	41.7	
June 10, 2000	Saturday	34	41.5	Washington, Newark
June 11, 2000	Sunday	34	44.6	Wilmington, Baltimore, Bristol, Newark
January 14, 2001	Sunday	26	45.5	
January 23, 2001	Tuesday	26	52.9	Baltimore, Washington, Wilmington
January 24, 2001	Wednesday	15	41.9	Baltimore, Washington, Wilmington
May 4, 2001	Friday	30	46.2	Baltimore, Washington
June 28, 2001	Thursday	34	42.9	Wilmington
June 29, 2001	Friday	34	49.2	Baltimore, Washington, Wilmington
June 30, 2001	Saturday	34	51.8	Wilmington, Newark
August 6, 2001	Monday	34	46.5	Baltimore, Washington, Wilmington, Newark
August 9, 2001	Thursday	34	50.4	Baltimore, Washington, Richmond, Wilmington
August 10, 2001	Friday	34	41.2	
November 18, 2001	Sunday	26	52.1	
June 9, 2002	Sunday	13	57.2	
July 2, 2002	Tuesday	34	42.8	Baltimore, Washington, Wilmington, Richmond
July 3, 2002	Wednesday	34	45.4	Baltimore, Washington, Richmond, Wilmington
July 9, 2002	Tuesday	34	44.0	Washington
July 18, 2002	Thursday	34	46.3	Baltimore, Washington, Richmond, Charlotte, Wilmington, Newark
July 19, 2002	Friday	34	58.5	Baltimore, Washington, Wilmington, Newark
August 13, 2002	Tuesday	34	40.9	Baltimore, Washington, Richmond, Wilmington

High PM concentrations were measured during all quarters at the Philadelphia sites with the maximum number of high days occurring during the summer months and the minimum number of high PM days occurring during the first quarter of the year. The Philadelphia area also measured high PM concentrations during the wintertime episode of January 23–24, 2001, and the summertime episodes discussed above: August 6–9, 2001, July 1–4, 2002, and July 18–19, 2002. Another widespread but short-term event occurred on August 13, 2002. During this period, a moderately strong upper-level ridge centered over the eastern states resulted in very light southwesterly winds aloft, and a moderately strong surface high-pressure system centered over Virginia. Minimum temperatures in the Philadelphia area were in the low 70's, while maximum temperatures were in the upper 90's. Hazy skies and fog were reported in the early morning hours at multiple sites throughout the region. These conditions led to high concentrations at sites extending from Richmond to Philadelphia. A USG day was observed in the Newark area on August 14. Meteorological conditions changed in the region on August 15, in advance of an approaching cold front, resulting in lower measured PM concentrations throughout the region.

This review of the meteorological conditions indicates that high PM concentrations occur under a variety of synoptic situations, but in general (and as indicated by the CART results) the majority of summertime events are associated with regional-scale build up and transport of PM, while the wintertime events seem to be driven by local meteorological conditions and can be isolated, depending upon the geographical extent of the PM conducive meteorological conditions. CART quite clearly distinguishes the winter- and summertime events and places a majority of these into two key bins. Other high PM days are placed in other high PM bins. CART thus appears to be able to distinguish and group the USG days quite effectively. Because of these differences, the categorical summaries should not be used to guide the forecasting, and instead the bin by bin characteristics must be considered.

4.10. Factors Influencing $PM_{2.5}$ Concentrations for Wilmington, DE

The area-wide maximum $PM_{2.5}$ for the Wilmington area was defined for this study as the maximum value over all of the sites listed as the local Wilmington sites in Table 2-1.

4.10.1. Summary of Observed $PM_{2.5}$ Data (1999–2002)

New Castle County in Delaware and Cecil County in Maryland provide data for Wilmington from six FRM monitors, plus three additional collocated monitors, each used as a back-up for the other monitor at its site. Two percent of these days are USG, with 19 occurring in the summer, six in the winter and one in the spring, as shown in Figure 4-17. Both summer and winter have fewer good than moderate days; Figure 4-18 shows peak monthly 90^{th} percentile values in June and January, and the lowest concentrations in March and September.

Figure 4-17. Distribution of 1999–2002 Days by Season and Severity: Wilmington

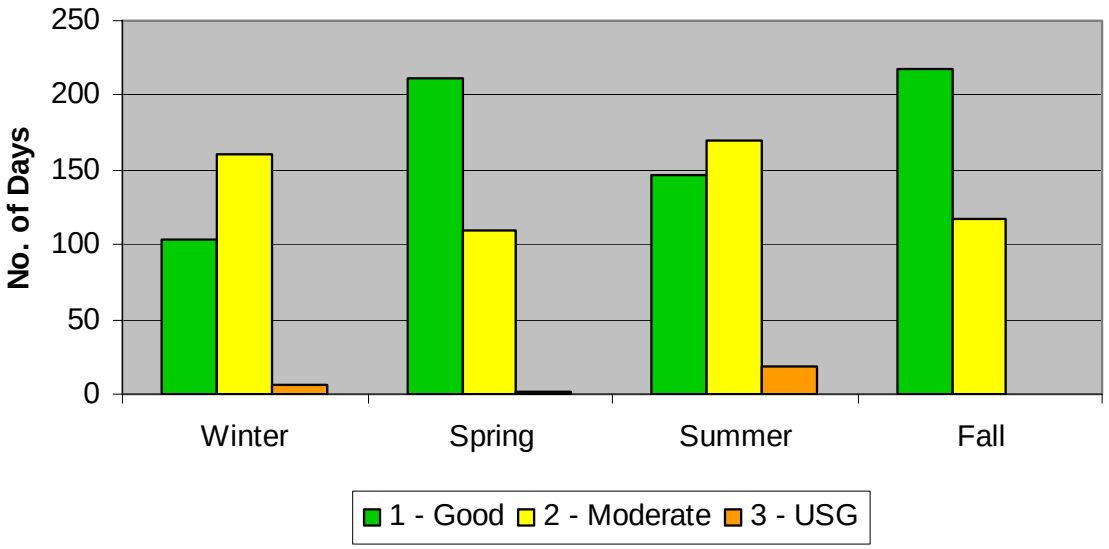
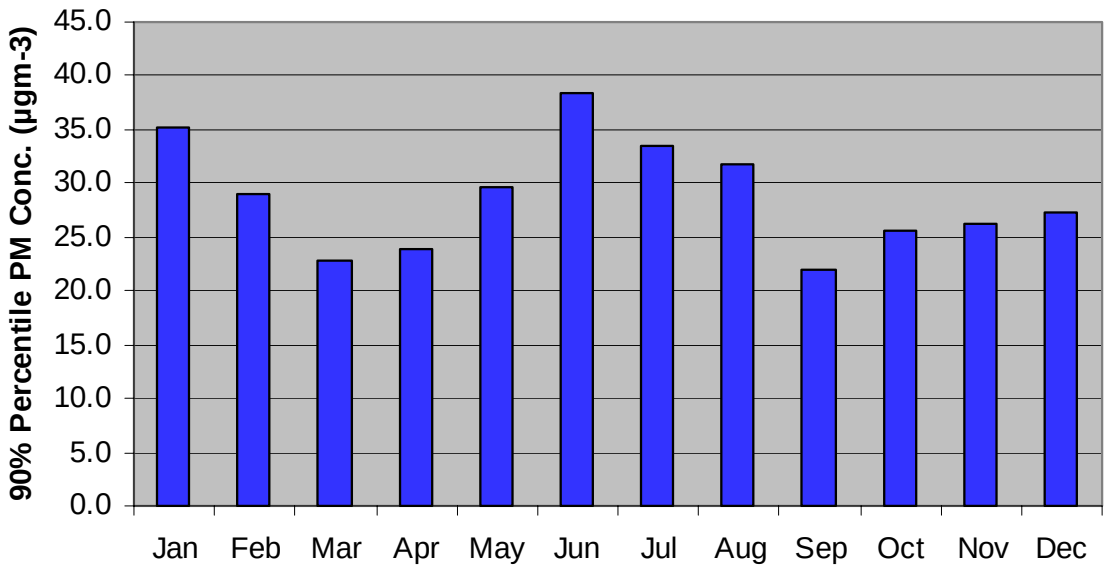


Figure 4-18. 90th Percentile Concentrations by Month (1999–2002): Wilmington



4.10.2. Meteorological Factors Influencing $PM_{2.5}$ Concentrations

The meteorological conditions associated with the different ranges of $PM_{2.5}$ concentration and specifically the highest PM days for the Wilmington area are discussed in this subsection.

Wind Patterns Associated with High $PM_{2.5}$

Plots comparing the frequency of wind directions and speeds for all days and for low, moderate, and high $PM_{2.5}$ days for the Wilmington area are presented in Appendix A. The wind information in

these plots is for the Dulles Airport (Sterling, VA) upper-air monitoring site. The plots use the same format and contain the same information as for the other areas (described earlier in this section).

The wind roses for Wilmington (Figures A-15 and A-16) are based on the Dulles Airport sounding data. The upper-level winds are predominately westerly to northwesterly for the low PM days at the time of the morning sounding, and southwesterly to northwesterly winds at the time of the evening sounding. For both sounding times, wind directions, on average, back to a more southwesterly direction for the moderate PM days, with lower wind speeds than for the lower PM days. For the highest PM days, wind speeds are much lower than for the other PM concentration levels and the wind directions generally range from southwesterly to northwesterly; at the time of the evening sounding many different directions are represented.

Categorical Summaries

A comparison of the meteorological characteristics for different ranges of $PM_{2.5}$ concentration in Table 4-23 provides a basis for further distinguishing days within the different categories based on the values of meteorological parameters. In preparing this table, we used the comprehensive meteorological and PM dataset compiled for the CART application. Key meteorological parameters, as used by CART to construct the classification tree, are shaded in this table so that we can focus on the differences in these key parameters as well the differences found throughout the dataset. Categories 1 to 3 represent the standard three ranges of 24-hour $PM_{2.5}$ concentration: <15.5, 15.5–40.5, and ≥40.5 μgm^{-3}.

Table 4-23. Summary of Mean Air Quality and Meteorological Parameters for Each CART Classification Category: Wilmington

	Category 1	Category 2	Category 3
$PM_{2.5}$ Parameters			
24-hour $PM_{2.5}$ for Wilmington (μgm^{-3})	10.3	22.9	47.4
Two-days-ago 24-hour $PM_{2.5}$ for New Castle (μgm^{-3})	15.2	17.1	21.5
Two-days-ago maximum 24-hour $PM_{2.5}$ for Washington, Baltimore, and Gettysburg (μgm^{-3})	17.0	18.8	27.1
Surface Meteorological Parameters			
Maximum surface temperature (°C)	17.0	19.2	26.5
Minimum surface temperature (°C)	8.0	9.3	16.0
Surface relative humidity (%)	65.7	69.8	70.1
Surface wind speed (ms^{-1})	3.7	2.6	2.1
Surface wind direction (degrees)	276	186	184
Number of six hour periods with precipitation (range is 1 to 4)	0.3	0.2	0.0
Upper-Air Meteorological Parameters (Dulles Airport)			
850 mb temperature (AM) (°C)	5.3	7.8	15.3
850 mb temperature (PM) (°C)	6.0	8.8	15.5

	Category 1	Category 2	Category 3
Temperature gradient (850 mb to surface; AM) (°C)	-3.3	-0.8	-1.8
Temperature gradient (900 mb to surface; AM) (°C)	-1.1	2.3	1.9
Temperature gradient (950 mb to surface; AM) (°C)	-0.3	3.2	3.9
24-hour difference in 700 mb geopotential height (m)	-1.4	-0.6	-17.3
Yesterday's 700 mb wind speed (PM) (ms^{-1})	14.2	12.7	6.8
Yesterday's 850 mb wind speed (PM) (ms^{-1})	10.5	9.0	6.1
850 mb wind speed (AM) (ms^{-1})	10.6	9.6	5.8
850 mb wind speed (PM) (ms^{-1})	10.1	9.6	6.9
Yesterday's 700 mb wind direction (PM) (degrees)	242	245	277
Yesterday's 850 mb wind direction (PM) (degrees)	242	229	257
850 mb wind direction (AM) (degrees)	253	237	247
850 mb wind direction (PM) (degrees)	246	227	243
Estimated cloud cover (range of 1 to 3)	1.9	1.8	1.7
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	2	2	3

Table 4-23 provides an overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for Wilmington. The results for Wilmington are very similar to those for Philadelphia.

High $PM_{2.5}$ in the Wilmington area is associated with relatively high $PM_{2.5}$ two-days prior—in both Wilmington and the Baltimore-Washington area. Thus, a regional day-to-day build up of $PM_{2.5}$ is indicated for high $PM_{2.5}$ days.

The surface meteorological parameters indicate a correlation between higher $PM_{2.5}$ concentrations and higher temperatures (primarily reflecting seasonal differences) and less precipitation. Surface wind directions tend toward southerly for the higher ranges of $PM_{2.5}$, compared to westerly for the lowest range of concentration. There is no clear tendency for relative humidity and surface wind speed.

The upper-air meteorological parameters (based on the Dulles Airport sounding) indicate that higher $PM_{2.5}$ concentrations occur with higher 850 mb temperatures and generally more stable (positive) lapse rates. The difference in geopotential height is much more negative for the higher PM days.

Considering the upper-air wind data, the higher PM days are characterized by lower wind speeds aloft. Winds aloft are, on average, southwesterly, for all three categories.

High PM is associated with slightly less cloud cover and tends to occur during the summer.

The input parameters that are most used by CART in the construction of the classification tree (either to define the splits/branching structure) or as surrogates to the primary variables in this regard are highlighted in the table. These include surface temperature, surface wind speed, 850 mb temperature, and 950 to surface temperature difference. All of these are also well correlated with the $PM_{2.5}$ concentration for the analysis day.

4.10.3. Characteristics of High $PM_{2.5}$ Events

The categorical summary table provides a general overview of how average conditions vary across (and potentially lead to) different $PM_{2.5}$ concentration levels for the Wilmington area. Within the high $PM_{2.5}$ categories, there are other key differences among the parameters that result in different types of high $PM_{2.5}$ events. We have used the CART results to examine these differences.

Only certain of the CART bins are frequently associated with $PM_{2.5}$ concentrations that are in the USG range or Category 3. Of these, we identified the bins with the most number of days as key bins. Table 4-24 considers the input parameter values for the key USG bins. For Wilmington there are two key bins that contain 15 and 5, respectively, of the 26 USG days.

Table 4-24. Summary of Mean Air Quality and Meteorological Parameters for Key USG CART Classification Bins: Wilmington

	Bin 29	Bin 2
Number of days	15	5
$PM_{2.5}$ Parameters		
24-hour $PM_{2.5}$ for Wilmington (μgm^{-3})	48.3	48.3
Two-days-ago 24-hour $PM_{2.5}$ for New Castle (μgm^{-3})	23.6	15.3
Two-days-ago maximum 24-hour $PM_{2.5}$ for Washington, Baltimore, and Gettysburg (μgm^{-3})	32.7	14.2
Surface Meteorological Parameters		
Maximum surface temperature (°C)	33.9	4.5
Minimum surface temperature (°C)	23.0	-5.2
Surface relative humidity (%)	64.2	81.2
Surface wind speed (ms^{-1})	2.7	0.6
Surface wind direction (degrees)	190	270
Number of six hour periods with precipitation (range is 1 to 4)	0.0	0.0
Upper-Air Meteorological Parameters (Dulles Airport)		
850 mb temperature (AM) (°C)	18.1	-0.9
850 mb temperature (PM) (°C)	18.9	0.1
Temperature gradient (850 mb to surface; AM) (°C)	-3.4	6.2
Temperature gradient (900 mb to surface; AM) (°C)	1.0	7.4
Temperature gradient (950 mb to surface; AM) (°C)	3.1	8.4
24-hour difference in 700 mb geopotential height (m)	-20.8	-36.2
Yesterday's 700 mb wind speed (PM) (ms^{-1})	5.7	8.8
Yesterday's 850 mb wind speed (PM) (ms^{-1})	5.3	7.3
850 mb wind speed (AM) (ms^{-1})	5.5	8.2
850 mb wind speed (PM) (ms^{-1})	6.8	9.1
Yesterday's 700 mb wind direction (PM) (degrees)	275	270
Yesterday's 850 mb wind direction (PM) (degrees)	252	270
850 mb wind direction (AM) (degrees)	248	270
850 mb wind direction (PM) (degrees)	246	252
Estimated cloud cover (range of 1 to 3)	1.7	1.5
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	3	1

The two key high PM bins represent winter and summer types of PM events.

Days within Bin 2 (containing a majority of winter time days) are associated with lower two-days-ago PM concentrations, yet similar concentrations, on average, on the analysis days, compared to days within Bin 29 (the summer time bin). Temperatures and surface wind speeds are much lower for days within Bin 2. Surface wind directions are also different for the two bins and are westerly for Bin 2 (winter) and southerly for Bin 29 (summer). The days within Bin 2 are also distinguished by more stable lapse rates and a deeper stable layer, than days within Bin 29—typical of wintertime conditions. Wind speeds aloft are greater for Bin 2 than for Bin 29, as expected during wintertime synoptic conditions.

Next we explore, the conditions associated with the USG events.

Data retrieval and availability for the Wilmington area were high for the period 1999–2002, and 26 USG days occurred during this period. The specific dates, including the observed $PM_{2.5}$ concentration (μgm^{-3}), are presented in Table 4-25.

Table 4-25. USG Days for Wilmington: 1999–2002

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG Day for Other Areas?
June 7, 1999	Monday	29	40.9	
June 8, 1999	Tuesday	29	45.5	
July 19, 1999	Monday	29	46.0	
July 24, 1999	Saturday	29	44.7	
July 31, 1999	Saturday	36	43.9	
January 1, 2000	Saturday	2	42.6	Charlotte, Washington
February 4, 2000	Friday	2	45.2	
February 10, 2000	Thursday	2	53.4	
June 11, 2000	Sunday	29	43.5	Baltimore, Bristol, Philadelphia, Newark,
January 13, 2001	Saturday	2	40.8	
January 23, 2001	Tuesday	2	59.6	Baltimore, Washington
January 24, 2001	Wednesday	4	58.5	Baltimore, Washington
May 19, 2001	Saturday	10	40.9	
June 14, 2001	Thursday	15	41.4	Newark
June 28, 2001	Thursday	29	41.9	Philadelphia
June 29, 2001	Friday	29	51.8	Baltimore, Washington, Philadelphia
June 30, 2001	Saturday	29	44.6	Newark, Philadelphia
August 6, 2001	Monday	29	44.9	Baltimore, Washington, Philadelphia, Newark
August 8, 2001	Wednesday	29	50.0	Baltimore, Washington, Richmond, Roanoke
August 9, 2001	Thursday	29	53.1	Baltimore, Washington, Richmond, Roanoke, Philadelphia

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG Day for Other Areas?
June 25, 2002	Tuesday	34	42.1	Baltimore, Washington, Philadelphia, Newark
July 2, 2002	Tuesday	29	57.8	Baltimore, Washington, Richmond, Philadelphia, Newark
July 3, 2002	Wednesday	29	46.1	Baltimore, Washington, Richmond, Philadelphia
July 18, 2002	Thursday	29	56.0	Baltimore, Washington, Richmond, Charlotte, Newark
July 19, 2002	Friday	29	57.6	Baltimore, Washington, Newark, Philadelphia, Wilmington
August 13, 2002	Tuesday	34	40.6	Baltimore, Washington, Richmond, Philadelphia

For the Wilmington area, the majority of high PM days were measured during the summer months, and no USG days occurred in the fourth quarter of the year (October-December). Due to its proximity to the Washington, Baltimore, and Philadelphia areas, the Wilmington area experiences similar meteorological conditions that lead to high PM concentrations. As noted above, the Wilmington area experienced multiple USG days during the January 23–24, 2001 wintertime episode, and during the widespread summertime episode periods of August 5–9, 2001 and July 1–4, 2002.

Another widespread episode that occurred in the MARAMA region was the July 17–19, 2002 period. Similar to the other summertime episodes, a strong upper-level ridge was centered over the Midwest during this period, with a surface high-pressure system centered over Georgia. Upper-level winds were light and southwesterly, while surface winds were light and variable. Maximum temperatures were in the upper 90's, while minimum temperatures were in the low 70's. Hazy skies and limited visibility were reported during the morning hours throughout the region. The meteorological conditions of this episode are very similar to those of the July 1–4, 2002 period. High PM was measured at six of the nine areas of interest, from Charlotte to the south extending to Newark to the north.

This review of the meteorological conditions indicates that high PM concentrations occur under a variety of synoptic situations. As for Philadelphia, CART distinguishes the winter- and summertime events and places a majority of these into two key bins. Other high PM days are placed in other high PM bins. CART thus appears to be able to distinguish and group the USG days quite effectively. Because of these differences, the categorical summaries should not be used to guide the forecasting, and instead the bin by bin characteristics must be considered.

4.11. Factors Influencing $PM_{2.5}$ Concentrations for Newark, NJ

The area-wide maximum $PM_{2.5}$ for the Newark/Elizabeth area was defined for this study as the maximum value over all of the sites listed as the local Newark sites in Table 2-1.

4.11.1. Summary of Observed $PM_{2.5}$ Data (1999–2002)

The data for Newark come from ten FRM monitors in three New Jersey counties in the Newark MSA: Essex, Middlesex, and Union. Two additional monitors are collocated with two others and

only used if data from the primary monitors are missing. Only 2.5 percent of these days are USG, all but three occur in the summer months. Figure 4-19 shows the distribution of days by season and severity. Winter has almost as many moderate days as good, though only one very high USG day; fall and spring have mostly good days. Figure 4-20 shows the 90th percentile concentrations by month, with the highest occurring in June and August, but second highest in October, followed by July, followed closely by January.

Figure 4-19. Distribution of 1999–2002 Days by Season and Severity: Newark

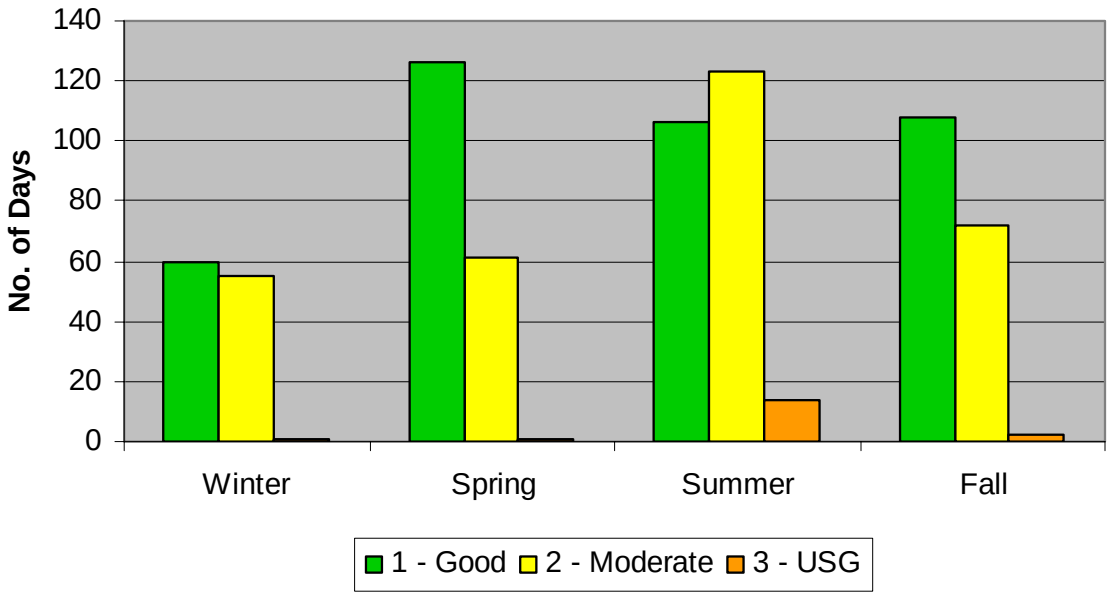
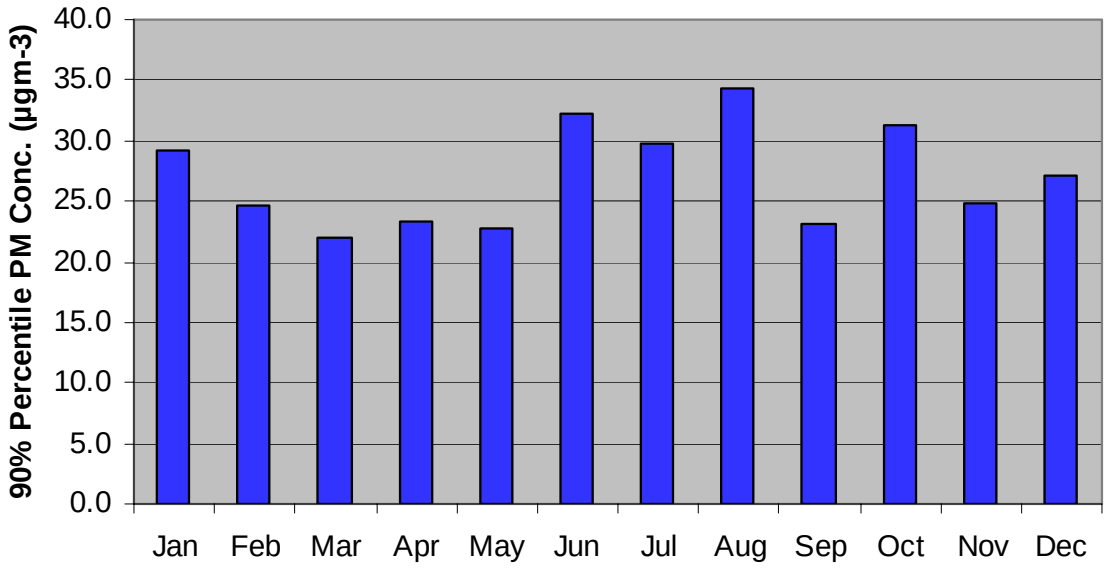


Figure 4-20. 90th Percentile Concentrations by Month (1999–2002): Newark



4.11.2. Meteorological Factors Influencing $PM_{2.5}$ Concentrations

The meteorological conditions associated with the different ranges of $PM_{2.5}$ concentration and specifically the highest PM days for the Newark area are discussed in this subsection.

Wind Patterns Associated with High $PM_{2.5}$

Plots comparing the frequency of wind directions and speeds for all days and for low, moderate, and high $PM_{2.5}$ days for the Newark area are presented in Appendix A. The wind information in these plots is for the Brookhaven (Long Island, NY) upper-air monitoring site. The plots use the same format and contain the same information as for the other areas (described earlier in this section).

The wind roses for Newark (Figures A-17 and A-18) are based on the Brookhaven sounding data. The upper-level winds are predominately west-southwesterly to northerly for the low PM days at the time of both the morning and evening soundings, northwesterly winds characterize the greatest number of days for the evening hour. For both sounding times, wind directions, on average, back to a more southwesterly direction for the moderate PM days, with lower wind speeds than for the lower PM days. The range in wind direction is southwesterly to northwesterly, and the greatest number of days with westerly winds. For the highest PM days, there is a further shift toward southwesterly and the predominant range in wind direction is southwesterly to westerly. Wind speeds are lower than for the other PM concentration levels.

Categorical Summaries

A comparison of the meteorological characteristics for different ranges of $PM_{2.5}$ concentration in Table 4-26 provides a basis for further distinguishing days within the different categories based on the values of meteorological parameters. In preparing this table, we used the comprehensive meteorological and PM dataset compiled for the CART application. Key meteorological parameters, as used by CART to construct the classification tree, are shaded in this table so that we can focus on the differences in these key parameters as well the differences found throughout the dataset. Categories 1 to 3 represent the standard three ranges of 24-hour $PM_{2.5}$ concentration: <15.5, 15.5–40.5, and ≥40.5 μgm^{-3}.

Table 4-26. Summary of Mean Air Quality and Meteorological Parameters for Each CART Classification Category: Newark

	Category 1	Category 2	Category 3
$PM_{2.5}$ Parameters			
24-hour $PM_{2.5}$ for Newark/Elizabeth (μgm^{-3})	9.6	23.4	45.4
Two-days-ago 24-hour $PM_{2.5}$ for Elizabeth (μgm^{-3})	14.9	15.8	22.8
Two-days-ago 24-hour $PM_{2.5}$ for Bethlehem (μgm^{-3})	13.2	15.2	22.0
Two-days-ago maximum 24-hour $PM_{2.5}$ for Camden and New Castle (μgm^{-3})	15.6	17.1	24.9
Surface Meteorological Parameters			
Maximum surface temperature (°C)	18.2	22.2	29.9
Minimum surface temperature (°C)	10.0	12.8	19.4
Surface relative humidity (%)	59.6	67.2	67.8

	Category 1	Category 2	Category 3
Surface wind speed (ms^{-1})	3.9	3.1	3.2
Surface wind direction (degrees)	268	176	169
Number of six hour periods with precipitation (range is 1 to 4)	0.2	0.2	0.2
Upper-Air Meteorological Parameters (Brookhaven)			
850 mb temperature (AM) (°C)	5.0	9.0	14.7
850 mb temperature (PM) (°C)	5.3	10.1	16.5
Temperature gradient (850 mb to surface; AM) (°C)	-4.4	-2.5	-3.9
Temperature gradient (900 mb to surface; AM) (°C)	-2.1	0.5	0.7
Temperature gradient (950 mb to surface; AM) (°C)	-1.2	1.7	2.5
24-hour difference in 700 mb geopotential height (m)	-0.3	4.1	1.1
Yesterday's 700 mb wind speed (PM) (ms^{-1})	14.3	12.7	10.8
Yesterday's 850 mb wind speed (PM) (ms^{-1})	11.1	9.3	7.9
850 mb wind speed (AM) (ms^{-1})	10.5	9.6	8.4
850 mb wind speed (PM) (ms^{-1})	10.2	10.3	10.0
Yesterday's 700 mb wind direction (PM) (degrees)	238	248	264
Yesterday's 850 mb wind direction (PM) (degrees)	252	243	250
850 mb wind direction (AM) (degrees)	257	228	236
850 mb wind direction (PM) (degrees)	260	237	204
Estimated cloud cover (range of 1 to 3)	1.8	1.9	1.6
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	2	2	3

Table 4-26 provides an overview of how average conditions for each classification category for Newark.

High $PM_{2.5}$ in the Newark area is associated with relatively high $PM_{2.5}$ two-days prior—in the Newark-Elizabeth area, as well as in the Camden-New Castle area and in Bethlehem, PA. Thus, a regional day-to-day build up of $PM_{2.5}$ is indicated for high $PM_{2.5}$ days.

The surface meteorological parameters indicate a correlation between higher $PM_{2.5}$ concentrations and higher temperatures (primarily reflecting seasonal differences), lower surface wind speeds, and higher relative humidity. Surface wind directions tend toward southerly, compared to westerly for the lowest PM range. There is no clear tendency with respect to wind speed or precipitation.

The upper-air meteorological parameters (based here on the Brookhaven, NY sounding) indicate that higher $PM_{2.5}$ concentrations occur with higher 850 mb temperatures. There is also some tendency for more stable (positive) lapse rates to be associated with higher $PM_{2.5}$ days. This is especially true for the 900 and 950 mb temperature differences.

Considering the upper-air winds, wind speeds are slightly lower aloft (especially for the analysis day); wind directions are similar for all three categories and, on average, southwesterly.

Finally, the cloud cover is less for the high PM days, the majority of which tend to occur, based on the season index, during the summer months.

The input parameters that are most used by CART in the construction of the classification tree (either to define the splits/branching structure) or as surrogates to the primary variables in this regard are highlighted in the table. These include surface temperature, relative humidity, 850 mb temperature, and 900 to surface temperature difference. All of these are also correlated with the $PM_{2.5}$ concentration for the analysis day. Newark is one of the few area for which relative humidity is a key CART parameter and varies regularly among the categories (increasing with increasing PM concentration).

4.11.3. Characteristics of High $PM_{2.5}$ Events

The categorical summary table provides a general overview of how average conditions vary across (and potentially lead to) different 8-hour ozone concentration levels for the Newark area. Within the high $PM_{2.5}$ categories, there are other key differences among the parameters that result in different types of high $PM_{2.5}$ events. We have used the CART results to examine these differences.

Only certain of the CART bins are frequently associated with $PM_{2.5}$ concentrations that are in the USG range or Category 3. Of these, we identified the bins with the most number of days as key bins. Table 4-27 considers the input parameter values for the key USG bins. For Newark there are two key bins and these contain 89 percent of the USG days.

Table 4-27. Summary of Mean Air Quality and Meteorological Parameters for Key USG CART Classification Bins: Newark

	Bin 34	Bin 13
Number of days	13	3
$PM_{2.5}$ Parameters		
24-hour $PM_{2.5}$ for Newark/Elizabeth (μgm^{-3})	43.0	46.9
Two-days-ago 24-hour $PM_{2.5}$ for Elizabeth (μgm^{-3})	21.8	22.8
Two-days-ago 24-hour $PM_{2.5}$ for Bethlehem (μgm^{-3})	22.8	18.1
Two-days-ago maximum 24-hour $PM_{2.5}$ for Camden and New Castle (μgm^{-3})	27.4	17.5
Surface Meteorological Parameters		
Maximum surface temperature (°C)	34.4	19.4
Minimum surface temperature (°C)	22.8	11.5
Surface relative humidity (%)	61.9	82.6
Surface wind speed (ms^{-1})	4.1	0.4
Surface wind direction (degrees)	175	0
Number of six hour periods with precipitation (range is 1 to 4)	0.2	0.0

	Bin 34	Bin 13
Upper-Air Meteorological Parameters (Brookhaven)		
850 mb temperature (AM) (°C)	17.0	9.4
850 mb temperature (PM) (°C)	18.8	10.1
Temperature gradient (850 mb to surface; AM) (°C)	-5.2	-1.8
Temperature gradient (900 mb to surface; AM) (°C)	0.2	1.3
Temperature gradient (950 mb to surface; AM) (°C)	2.0	2.7
24-hour difference in 700 mb geopotential height (m)	3.1	0.0
Yesterday's 700 mb wind speed (PM) (ms^{-1})	10.4	9.9
Yesterday's 850 mb wind speed (PM) (ms^{-1})	8.7	6.7
850 mb wind speed (AM) (ms^{-1})	9.4	4.1
850 mb wind speed (PM) (ms^{-1})	9.9	7.9
Yesterday's 700 mb wind direction (PM) (degrees)	270	180
Yesterday's 850 mb wind direction (PM) (degrees)	264	207
850 mb wind direction (AM) (degrees)	252	207
850 mb wind direction (PM) (degrees)	212	180
Estimated cloud cover (range of 1 to 3)	1.6	1.7
Seasonal indicator (1 = winter, 2 = transitional, 3 = summer)	3	2

The two key high PM bins represent transitional period and summer of PM events.

Days within Bin 3 (containing a majority of transitional period days) are associated with lower two-days-ago PM concentrations at the upwind sites but higher values at the local site. The PM concentrations are also higher, on average, on the analysis days, compared to days within Bin 34 (the summertime bin). Temperatures and surface wind speeds are much lower for days within Bin 3. Surface wind directions are also different for the two bins and are northerly for Bin 2 (winter) and southerly for Bin 29 (summer). The days within Bin 2 are also distinguished by slightly more stable lapse rates than days within Bin 34. Wind speeds aloft are lower for Bin 3 than for Bin 34. Wind directions aloft are southerly, on average, for days in Bin 3, and westerly to southwesterly for days within Bin 34. These differences are similar to the winter/summer differences for the key bins for the Philadelphia and Wilmington area, but less dramatic.

Next, we examine the conditions associated with each day or episode.

Data retrieval and availability for the Newark area were moderate for the period 1999-2002, and 18 USG days occurred during this period. The specific dates, including the observed $PM_{2.5}$ concentration (μgm^{-3}), are listed in Table 4-28.

Table 4-28. USG Days for Newark: 1999–2002

Date	Day of Week	CART bin	$PM_{2.5}$ (μgm^{-3})	USG day for other areas?
June 2, 2000	Friday	34	41.6	
June 10, 2000	Saturday	34	45.0	Philadelphia
June 11, 2000	Sunday	34	41.6	Baltimore, Bristol, Wilmington, Philadelphia
October 26, 2000	Thursday	13	54.6	Baltimore, Washington
October 27, 2000	Friday	30	77.7	Washington, Bristol
December 11, 2000	Monday	13	44.9	
June 14, 2001	Thursday	34	43.4	Wilmington
June 30, 2001	Saturday	34	46.4	Wilmington, Philadelphia
August 6, 2001	Monday	34	41.0	Baltimore, Washington, Wilmington
August 10, 2001	Friday	34	42.4	
March 15, 2002	Friday	30	40.6	
June 11, 2002	Tuesday	34	42.5	
June 26, 2002	Wednesday	34	42.1	
July 2, 2002	Tuesday	34	41.8	Baltimore, Washington, Richmond, Wilmington, Philadelphia
July 18, 2002	Thursday	34	43.4	Baltimore, Washington, Richmond, Charlotte, Wilmington, Philadelphia
July 19, 2002	Friday	13	41.7	Baltimore, Washington, Wilmington, Philadelphia
August 13, 2002	Tuesday	34	43.9	Baltimore, Washington, Richmond, Wilmington
August 14, 2002	Wednesday	34	44.0	

Although data retrieval for the Newark area was less than that for Philadelphia, Wilmington, Baltimore, and Washington, a number of high $PM_{2.5}$ events were measured during the 1999–2002 period.

The Newark area experienced high PM concentrations during the summertime episodes discussed above: August 6–10, 2001, July 1–4, 2002, July 17–19, 2002, and August 13–14, 2002.

Very high PM was measured during one fall episode in the MARAMA region during the period October 24–27, 2000. During this period, the area was influenced by a moderately strong upper-level ridge centered over the Midwest. A strong, persistent surface high-pressure system was centered directly over the mid-Atlantic states and gradually weakened and moved northeastward by the end of the period. Low temperatures were in the mid 50's, with highs in the low 70's. Partly cloudy skies and fog were reported in the early morning hours throughout the region. Surface winds were very light, reflecting stagnation conditions, allowing for a buildup of PM concentrations throughout the region. In addition to the Newark area, USG level concentrations were measured at Baltimore and Washington on October 26 and at Baltimore, Washington, and Bristol on October 27.

This review of the meteorological conditions indicates the high PM concentration occur under a variety of synoptic situations, that vary according to season. There are two key USG bins for Newark and these represent summertime and transitional-period conditions. Other high $PM_{2.5}$ days are placed in other high PM bins. CART thus appears to be able to distinguish and group the USG days quite effectively. Because of these differences, the categorical summaries should not be used to guide the forecasting, and instead the bin by bin characteristics must be considered.

This page intentionally left blank.

5. $PM_{2.5}$ Forecasting Tools

CART-based forecasting tools were developed for each of the areas of interest. The forecasting algorithms were based on the CART binary decision trees and supporting information. Each tool consists of an interface for the entry of observed and forecasted data and other parameters, the forecasting algorithms and supporting calculations for one or more areas, and several options for the display, summary, and storage/archival of the input parameters and the forecast results.

Tools were developed for each of three sets of CART results. These included a draft version of the "operational" tools (based on the Regional 2 CART analysis), a final version of the "operational" tools (based on the Regional 3 CART results), and a "research" version of the tools (based on the Research CART results).

For each of the three sets of CART results, four tools were developed for: 1) Charlotte; 2) Bristol, Roanoke, and Richmond; 3) Baltimore, Washington, Philadelphia, and Wilmington; and 4) Newark. When multiple areas are included, the user must select the forecast area and forecasts are prepared one area at a time. The combined tools facilitate the preparation of forecasts for multiple areas (using only one program) and also allow the upper-air data that is input for one area to be used in preparing the forecast for another without reentry.

The forecast tools are described in this section. An evaluation of the tools using real-time and historical data is also presented.

5.1. Description of the CART-Based Forecasting Tools

The following description of the CART-based $PM_{2.5}$ forecasting tools includes an overview of the concepts, input requirements, features, and output summaries.

5.1.1. Conceptual Overview

By providing detailed information about the classification of historical days into bins with different $PM_{2.5}$ concentration ranges based on the values of related meteorological and air quality parameters, the CART trees provide a basis for similarly classifying future days based on the observed and predicted values of these same parameters. Specifically, the observed data and forecast parameters corresponding to a future day are compared with the decision points that define the CART tree and assigned to one of the classification bins. The path taken through the CART tree and the resulting classification is determined by the values of the observed data and forecast parameters and the binary splits that comprise the classification tree. The forecasted $PM_{2.5}$ concentration is assigned the value of the CART bin into which the day is classified.

This approach to forecasting has several attributes. Compared to simple regression techniques, the use a CART-based forecasting algorithm accommodates the possibility that different meteorological conditions can lead to the same or similar $PM_{2.5}$ concentrations and, most importantly, that there may be multiple pathways to high $PM_{2.5}$. The parameter and parameter values associated with the CART classification tree provide information on the relative importance of the various air quality and meteorological parameters to the air quality conditions as represented by the dependent variable. Thus the CART technique offers additional physical insight into phenomena being studied. By segregating the data values into the classification bins, CART also provides information regarding the frequency of occurrence of the conditions associated with each classification category. In this manner, the likely recurrence rate for a particular type of day and the associated prevalent conditions are obtained.

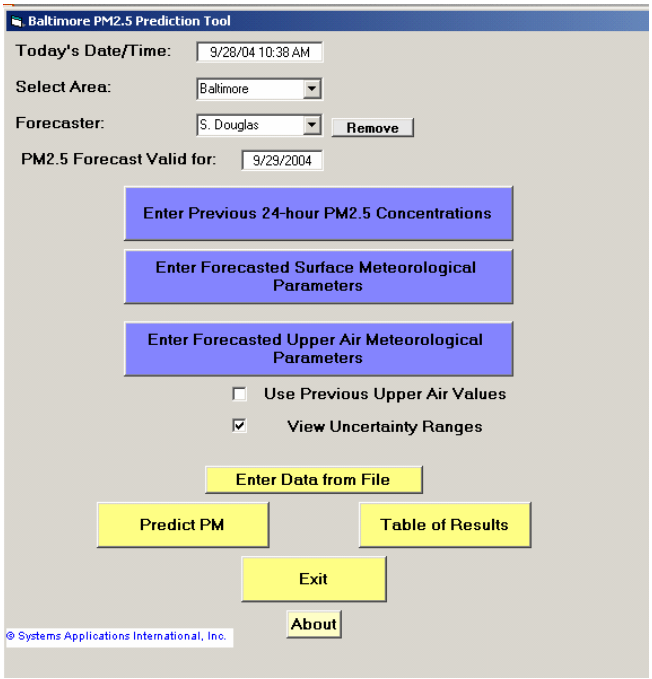
Two key assumptions come into play in the use of the CART result in this way. First, we assume that the relationships identified by CART and defined by the classification tree are physically meaningful. Our review and quality assurance of the CART outputs helps to ensure this, but it is important to keep in mind that CART is a statistical tool and not all of the identified statistical relationships can be confirmed to be physically meaningful (in part due to the complex nature of $PM_{2.5}$ formation and transport, and in part due to the complexity of the CART results). Next we assume that the CART application is complete with respect to representation of both the full range of different $PM_{2.5}$ regimes as well as the full set of input parameters needed to characterize the different regimes. Use of a limited dataset (in this case, a three- to four-year dataset) affects our ability to represent the range of regimes. The robustness of the input parameters is limited by the number and type of measurements, the spatial and temporal resolution of the "data", and the quality of the "data" in both the historical and forecast modes.

5.1.2. Input Requirements

In discussing the input requirements, we begin with some basic information that is either supplied by the tool or must be supplied by the forecaster. Basic forecast elements such as the date and time at which the forecast is made "Today's Date and Time" and the date for which the forecast is valid "PM2.5 Forecast Valid For" are supplied automatically by the tool. The forecast valid date is automatically set to tomorrow's date but can be changed by the forecaster. The user may enter his or her name "Forecaster" and for the multiple-area tools, must select an area "Select Area." The initial input screen for an example application for Baltimore is displayed in Figure 5-1.

All other input parameters are described in some detail in Section 2 of this report and in more operational terms in the next few subsections.

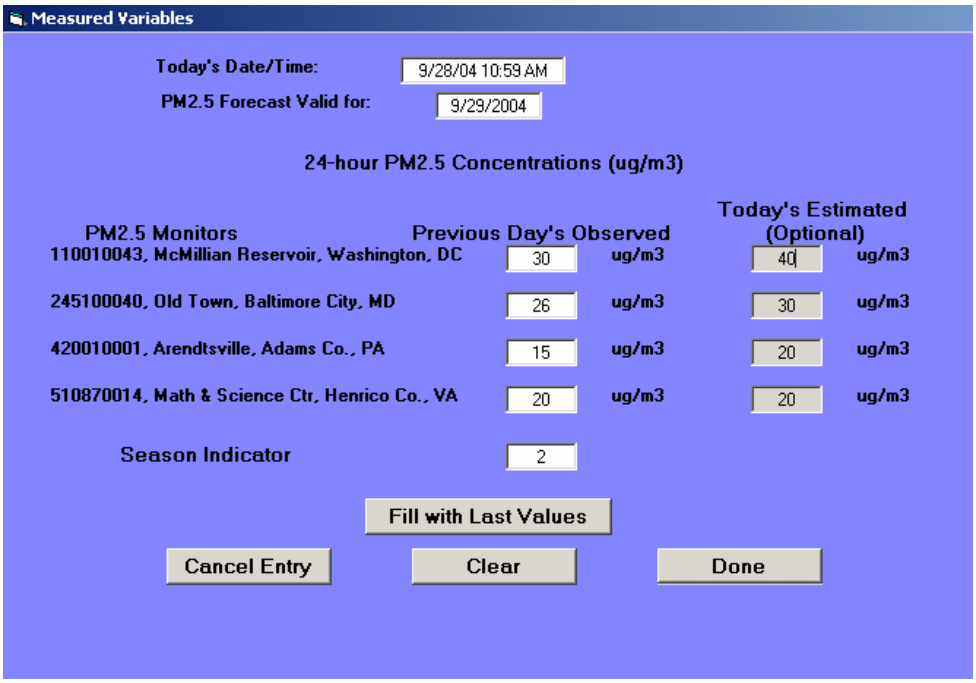
Figure 5-1. Initial Input Screen for the PM2.5 Forecasting Tool: Example for Baltimore



PM and Other Input Parameters

The first input screen is for entry of the "Previous 24-hour PM2.5 Concentrations". An example of this input screen is given in Figure 5-2. These inputs must be entered by hand. The user must provide the observed values for $PM_{2.5}$ for each site listed, for two days prior to the forecast day. We have also included a second column for estimated $PM_{2.5}$ values for the day prior to the forecast day. This information is required for the research version of the tool and is optional for the operational version of the tool. We suggest that consideration and entry of the one-day prior values may help with the review, interpretation and subsequent use of the CART-based forecast results.

Figure 5-2. Example Input Screen for PM2.5 Data



It is important to keep in mind that the CART-based forecasting tools were developed using PM2.5 data from FRM measurement systems—as they are expected to provide the most consistent and accurate concentration values. However, because they are collected using filters, FRM data are typically not available until several weeks after the sampling date. Thus, forecasters must rely on continuous measurements of $PM_{2.5}$ (which are available on a real time basis) to provide information about prior day PM levels at local and upwind sites and to support the forecasting. There are several different types of instruments used to collect continuous data, and these do not always agree with the FRM measurements. The level of disagreement varies from site to site, and typically from season to season (with temperature and humidity), as discussed in some detail by Gillespie et al. (2004). The use of the real-time data from continuous measurement systems may be different enough from the FRM data under some circumstances to cause an erroneous forecast. For most areas, prior day PM2.5 concentrations were important to the CART analysis and thus to the forecasts - increasing the possibility that differences in the data types could contribute to forecast errors.

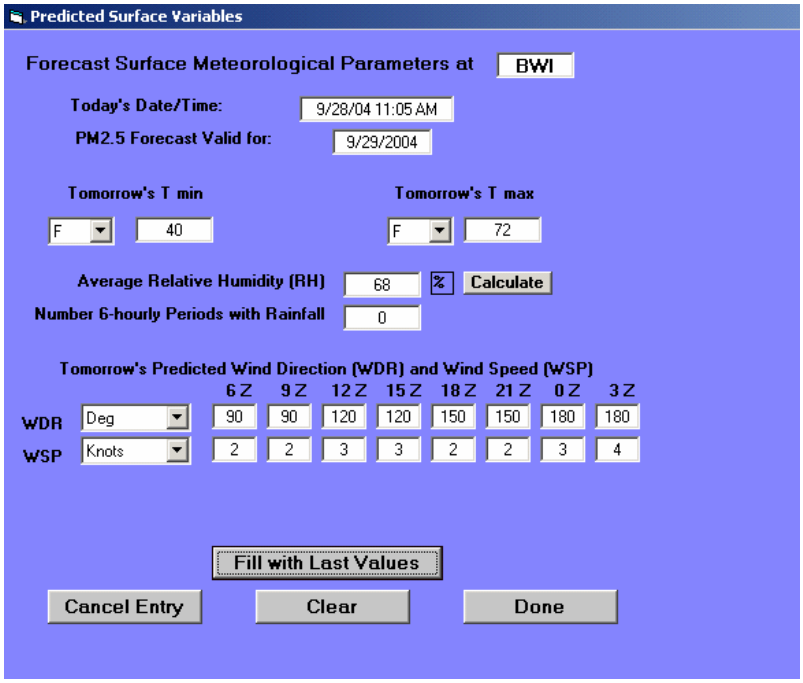
In specifying the prior-day $PM_{2.5}$, the forecaster should consider the whether the TEOM (or other real-time) data should be adjusted to account for differences between these data and the FRM data (as used in the underlying CART analysis).

The user must also specify the seasonal period of the forecast day. To account for seasonal variations in vegetative cover, there are three periods to choose from. The winter period includes November, December, January, February, and March. The transitional period includes April, May, September, and October. The summer period includes June, July, and August. This is an input rather than automatically generated to allow the user to choose different periods than appropriate for the date, for example, during transitional times or to accommodate unusual meteorological conditions such as drought.

Surface Meteorological Parameters

The second input screen is for entry of "Forecasted Surface Meteorological Parameters". An example of this input screen is given in Figure 5-3. These inputs may either be entered by hand or using the automated data entry feature, as discussed in the next section on features of the tool. The surface meteorological inputs are listed and described in Table 2-5. Care should be taken to specify the correct units for each parameter, as appropriate. Relative humidity (daily average) can either be entered directly or calculated based on 3-hourly values of temperature and dew-point temperature. Note that the typical forecast products provide the surface values at three-hourly intervals. The expected meteorological monitoring site will appear at the top of the screen. For Washington, D.C., surface winds from Dulles Airport (IAD) are recommended.

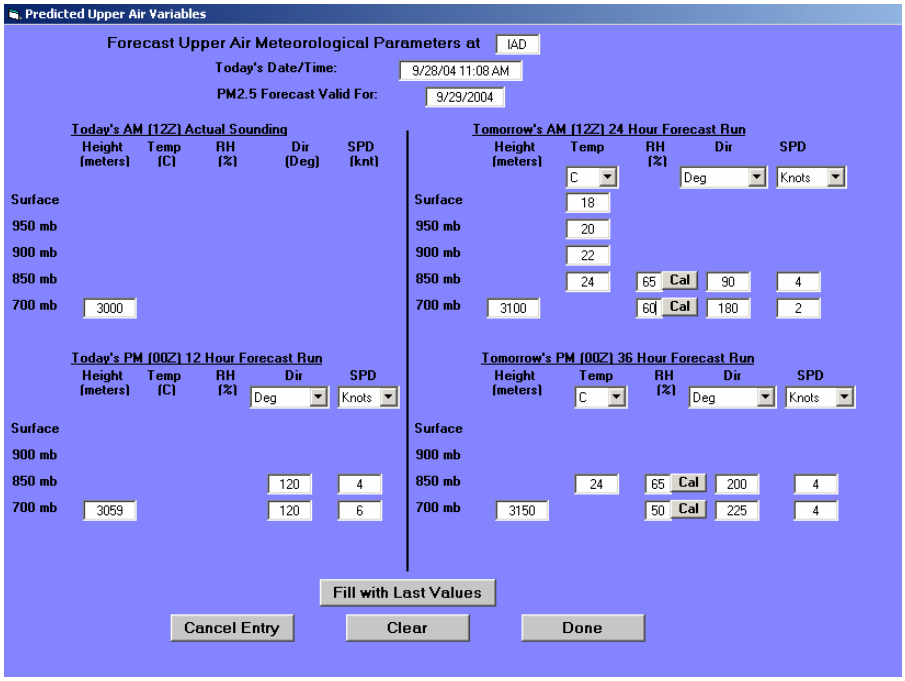
Figure 5-3. Example Input Screen for Surface Meteorological Data



Upper-Air Meteorological Parameters

The third input screen is for entry of "Forecasted Upper-Air Meteorological Parameters". An example of this input screen is given in Figure 5-4. These inputs may either be entered by hand or using the automated data entry feature, as discussed in the next section on features of the tool. The upper-air meteorological inputs are listed and described in Table 2-6. Care should be taken to specify the correct units for each parameter, as appropriate. Relative humidity can either be entered directly or calculated based on predicted values of temperature and dew-point temperature. The expected meteorological monitoring site will appear at the top of the screen. The entries are organized in chronological order and then by level (with increasing vertical height) for each required variable.

Figure 5-4. Example Input Screen for Upper-Air Meteorological Data



5.1.3. Features

Automated Data Entry

The surface and upper-air meteorological inputs can be entered by hand or can be read in from external data files. For the MARAMA project, surface and upper-air meteorological inputs are prepared on a daily basis by meteorologists from the Pennsylvania Department of Environmental Protection (DEP) and posted to a MARAMA forecaster's web site (S. Nolan, personal communication). There are currently three options for the obtaining the surface input parameters from the web site. The parameters are derived from the output for three different models including the NWS ETA model, the Global Forecast Systems (GFS) model, and the Nested Grid Model (NGM). The upper-air parameters are currently available for the ETA model only. The parameters, levels, and units are designed to match those required by the forecast tools.

Other Input Related Features

For tools that contain multiple areas, the upper-air data for a given upper-air monitoring site need not be entered twice. Instead, the user can check the box on the first form that is labeled "Previous Upper Air Values" to use the last entered data for the assigned upper-air site.

At the bottom of each data input screen is a box labeled "Fill with Last Values." This option allows the user to quickly make changes to one or more of the previously entered input parameters. This feature allows the forecaster to explore how small changes in one or more of the input parameters affect the forecast result.

Once the data for each category have been entered the box on the first screen will change color. When the inputs for all three categories have been provided by the user, the tool is ready to prepare a forecast.

Forecast Probabilities

The CART-based probabilities associated with the forecast bins are reported as part of the forecast. These characterize the probability for a day within the bin to belong to the classification category to which that bin is assigned or to belong to another classification category. This takes into account the number of days within the bin, weighted by the observed data distribution and the misclassification costs.

Forecast Range

The forecasting accuracy will depend upon the accuracy of the input data and, in particular, the meteorological forecasts. Errors or uncertainties in the meteorological forecasts will translate into errors or uncertainties in the $PM_{2.5}$ forecasts. To address the issue of uncertainty in the meteorological input data and its effect on the $PM_{2.5}$ forecast, we have included an uncertainty feature. This feature can be selected by checking the "View Uncertainty Ranges" box on the first form.

The uncertainty feature allows the user to run the forecast and obtain results for two alternate forecast scenarios. For the "High" forecast, the parameters are adjusted to be generally more conducive to higher $PM_{2.5}$ concentrations as follows:

- Wind speeds reduced by 0.5 ms^{-1}
- Temperatures increased by 1.5°C
- Temperature differences (stability parameters) increased by 0.5°C.

For the "Low" forecast, the parameters are adjusted to be generally more conducive to higher $PM_{2.5}$ concentrations as follows:

- Wind speeds increased by 0.5 ms^{-1}
- Temperatures lowered by 1.5°C
- Temperature differences (stability parameters) decreased by 0.5°C.

The objective of this feature is to allow the user to assess the potential uncertainty of the forecast due to uncertainties in the meteorological forecasts and rounding of the meteorological

forecast data. A result matching the main prediction result indicates that small uncertainties in the meteorological forecast will not affect the predicted $PM_{2.5}$ level, but a change in either low or high $PM_{2.5}$ colors and bins with respect to the main prediction indicates that the prediction is subject to uncertainty. This feature is intended to provide perspective regarding the sensitivity of the forecast to small errors or uncertainties in the meteorological forecasts.

5.1.4. Outputs

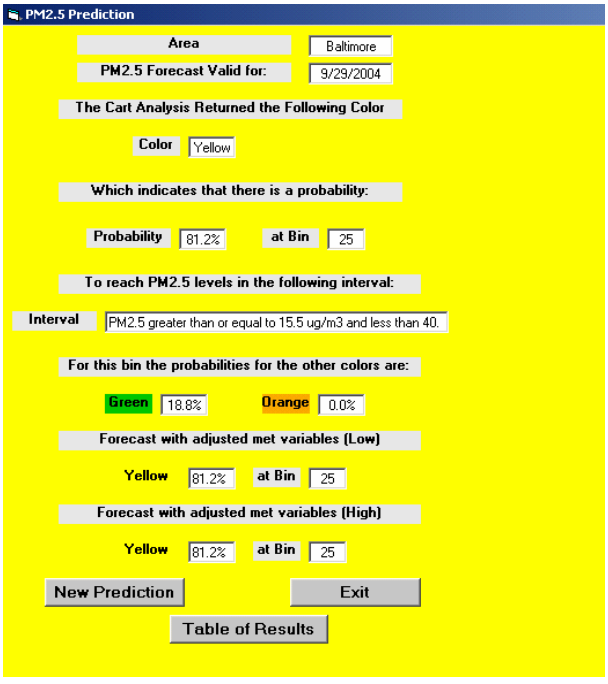
Once the data are entered, select "Predict PM" to obtain a forecast for 24-hour $PM_{2.5}$ for tomorrow (using the color-based air quality index (AQI)). The inputs and results will be presented on the screen and also summarized in a table. These tables can be used to check the inputs and to record the inputs and outputs.

Forecast Result

A primary output of the tool is the CART bin number (the bin into which the forecast day was placed) and the corresponding $PM_{2.5}$ concentration range for that bin. The forecast colors and ranges are as follows: Green (less than 15.5 μgm^{-3}), Yellow (15.5 to less than 40.5 μgm^{-3}), and Orange (greater than or equal to 40.5 μgm^{-3}). These correspond to "Good", "Moderate", and "USG", forecasts. The colors and ranges are indicated in the output. An example forecast result in given in Figure 5-5.

The forecast also includes the probabilities associated with the bin and, if requested by the user, the bin number and corresponding $PM_{2.5}$ range for the high and low forecasts.

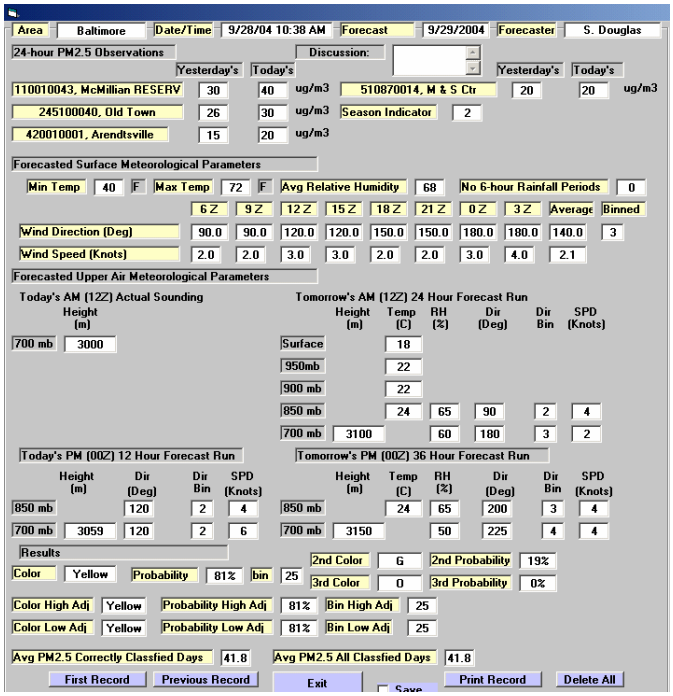
Figure 5-5. Example Forecast Result Screen



Summary of Results Table

The results table summarizes the various input parameters as well as the forecast result and supporting information. An example summary of results table is given in Figure 5-6. In addition, the average values for (1) all correctly classified days within the bin and (2) all days within the bin are given in the summary table. These values are based on the historical days that were place in that bin, and may provide some additional perspective to the forecast range. Space is provided for the user to enter forecast notes into the summary of results table.

Figure 5-6. Example Summary of Results Table



Archiving the Outputs

The tabular summaries are automatically saved within the database tool for each forecast. From the tool, the user may view the summary tables for previous forecasts, by paging through the archive of summary tables.

Selected tabular summaries can be exported to an Excel file, by checking the “Save” box for each table of information that is to be exported and then clicking on “Save As.” Only those tables that are checked will be exported.

All summary tables can be deleted using the “Delete All” button. This will clear all outputs from the tool.

5.2. Evaluation

In this section, we describe the methods and results of the evaluation of the draft version of the operational tools that was performed to air their refinement and the subsequent development of

the final versions of the tools. The evaluation discussed here concerns the tool's predictive accuracy, which depends on the CART tree itself, and not the user interface or other elements not based on CART but also evaluated and improved throughout the project.

5.2.1. Real-time Evaluation

Meteorologists in six of the nine MARAMA areas tested the draft versions of the operational $PM_{2.5}$ forecasting tools during February and March of 2004. For as many days as possible, each participant input the measured and forecasted meteorological and air quality data required by the tool to predict the next day's $PM_{2.5}$ level. "Good" days have maximum $PM_{2.5}$ concentrations less than 15.5 μgm^{-3}, "moderate" days have concentrations greater than or equal to 15.5 and less than 40.5 μgm^{-3}, and "USG" days have concentrations of 40.5 μgm^{-3} or above. The CART predictions were recorded and sent to a single person, who consolidated the results for each site and compared these to $PM_{2.5}$ observations from continuous monitors within each area. In this comparison, one continuous monitor was selected to represent each area: Garinger for Charlotte, Math & Science Center for Richmond, McMillan Reservoir for Washington, D.C., Old Town for Baltimore, Camden for Philadelphia, and MLK for Wilmington. Later, the CART predictions were also compared to quality-assured FRM data, which is compiled some time later than the continuous data. The data used for the evaluation are area-wide maximums over several FRM monitors within the area, as similar as possible to the area-wide maximums used to characterize each site during the pre-tool CART analysis. Thus this second comparison is closest to evaluating what the CART trees were originally designed to predict. At the time of the study, four MARAMA areas had sufficient first-quarter 2004 FRM data to undergo this second evaluation: Baltimore, Charlotte, Richmond, and Wilmington.

Several metrics were used to compare the $PM_{2.5}$ forecasting tool predictions to the observed continuous or FRM data. A simple matrix tallied how many days observed in each $PM_{2.5}$ category were forecast into each level. Accuracy, false alarm rate, probability of detection, critical success index (threat score), and bias statistics were derived from this information. The false alarm rate equals the percent of predicted USG days that did not turn out to be USG. The probability of detection equals the percent of observed USG days that were predicted to be USG. The critical success index is the number of successfully predicted USG days divided by the sum of false USG predictions and unpredicted USG days, and the bias is the ratio of number of predicted USG days over the number of observed USG days. In practice, these last three metrics were rarely of use since USG observations only occurred in two instances, both for Baltimore using FRM data. Therefore the false alarm rate and accuracy were the most informative measures, and the latter for the most part measured the tools' ability to tease out Good and Moderate days. A tool for the calculation of these metrics was provided by M. Seybold from the Maryland Department of the Environment (MDE).

The metrics described above were applied to the six areas in two different ways. The first, "strict" evaluation is a straightforward comparison of predicted and observed PM levels using the metrics described above. The second, "fuzzy-border" evaluation represents a best-case scenario by counting predictions as correct if the observed PM concentration fell within a designated border zone between the observed and predicted PM levels. For example, a Moderate prediction would be counted as correct even if the observed value is 14 $\mu g/m^3$, a little below the cut-off of 15.5 $\mu g/m^3$. The border zones are defined as follows: Good and Moderate predictions are both correct for concentrations greater than or equal to 13.5 $\mu g/m^3$ and less than or equal to 17.5 $\mu g/m^3$; Moderate and USG predictions are both correct for concentrations greater than or equal to 36.5 $\mu g/m^3$ and less than or equal to 44.5 $\mu g/m^3$.

Results of the "strict" and "fuzzy-border" evaluations are described below for the evaluation with continuous $PM_{2.5}$ observations and for the evaluation with FRM observations.

Real-time Evaluation Using Continuous PM Observations

Table 5-1 below provides statistics for the six sites evaluated for their ability to predict $PM_{2.5}$ levels indicated by a local continuous monitor. In addition to the accuracy and false alarm statistics, the table lists the number of days evaluated and the percentage of these days with "Good" PM levels. No USG days were observed, so the remainder of the days are all Moderate. Because no USG days were observed, the bias, critical success index, and probability of detection metrics were not included in the chart.

Table-5-1. Evaluation Metrics for PM Tools Using Continuous PM Observations

MARAMA Area—Monitor	No. Days	% Good Days	Accuracy (Strict)	False Alarm Rate (Strict)	Accuracy (Fuzzy)	False Alarm Rate (Fuzzy)
Charlotte, NC—Garinger	35	80%	66%	na*	71%	na*
Richmond, VA—Math & Sci. Ctr	38	79%	74%	na*	87%	na*
Washington, DC—McMillan	32	84%	75%	100%	88%	100%
Baltimore, MD—Oldtown	34	59%	68%	na*	77%	na*
Philadelphia, PA—Camden	29	59%	55%	na*	75%	na*
Wilmington, DE—MLK	37	35%	73%	na*	81%	na*

**No USG days predicted or observed*

Prediction accuracy ranges from 55 to 75 percent under the strict evaluation, and from 75 to 88 percent under the fuzzy evaluation. It is important to keep in mind in reviewing these percentages that all of the days exhibited low (good) or moderate $PM_{2.5}$ levels. One way to evaluate the predictive ability of the tools is to compare the accuracy to the accuracy if one had simply predicted all Good days (or all Moderate, in the case of Wilmington). Compared to the results using only one consistent forecast, the forecasting tool for Wilmington does a good job of predicting PM levels, the Baltimore tool does fairly well, the Philadelphia, Richmond, and Washington tools do barely well, and Charlotte does not do well at all. But this is a naive measure since prediction of very high PM days most concerns the forecaster, rather than the distinction between Good and Moderate. No high PM days were observed at the continuous monitors in February and March of 2004, fortunately for air quality but unfortunately for tool evaluation.

Real-time Evaluation Using FRM PM Observations

Table 5-2 below provides statistics for the four sites evaluated for their ability to predict $PM_{2.5}$ levels indicated by the maximum $PM_{2.5}$ concentration over several FRM monitors selected from the area. As in the previous section, the table gives the percentage of Good days according to the FRM data. No USG days were predicted for these four areas during the period, so false alarm rates are not shown.

Observed USG days appeared only for Baltimore; these two days were classified as Moderate by the forecasting tool so the table shows 0 percent as the detection probability; the critical success index and bias for Baltimore are also zero under strict evaluation, and nonexistent under fuzzy evaluation as the two USG days were below 44 µg/m^3 and therefore are almost Moderate. Fifty-nine percent of Baltimore's days had Moderate PM levels, according to the FRM data.

Table 5-2. Evaluation Metrics for PM Tools Using FRM PM Observations

MARAMA Area	No. Days	% Good Days	Accuracy (Strict)	Detection Prb. (Strict)	Accuracy (Fuzzy)	Detection Prb. (Fuzzy)
Charlotte, NC	35	60%	57%	na*	69%	na*
Richmond, VA	38	78%	78%	na*	89%	na*
Baltimore, MD	34	35%	50%	0%	65%	na*
Wilmington, DE	37	52%	84%	na*	95%	na*

**No USG days predicted or observed*

The $PM_{2.5}$ forecasting tools for Richmond and Wilmington appear to do a genuinely good job, although during this period their ability to predict high PM days remained untested. Agreement with the FRM data is better than with the continuous data in both cases. Forecasting ability is fair for Baltimore and Charlotte, regardless whether strict or fuzzy-border evaluations are considered. Agreement with the FRM data is worse than with the continuous data in both cases. The greatest changes in performance when the FRM data area used appear for Wilmington and Baltimore. The $PM_{2.5}$ levels for Wilmington tended to be lower according to the FRM monitors than according to the continuous monitor, whereas the opposite is true for Baltimore; this suggests uncertainty in actual PM concentrations, something to consider while evaluating $PM_{2.5}$ forecasting tools in real time.

5.2.2. Historical Period Evaluation

Historical data enabled evaluation of the forecasting tools for all nine areas. The same "strict" and "fuzzy-border" procedures described above were applied to the period of June through August, 2003, by running the data for these months through the classification tree using CART software rather than the forecasting tool. The summer 2003 data were prepared for CART in almost the same way the 1999–2002 data were prepared in creation of the original CART trees. The only difference was that some alternate FRM sites were used for the 2003 dependent value data, in instances where the original FRM monitor was shut down and replaced with another. So the observed data in this evaluation are more like the FRM data than the continuous data of the real-time comparisons described above; the 2003 and 2004 FRM-based PM datasets mirror as closely as possible the original 1999–2002 $PM_{2.5}$ data classified by CART.

The advantage of this method is that one can swiftly evaluate the tree using many datapoints (around ninety days for most of the areas). On the other hand, the evaluation is not exactly the same as if it were conducted using the PM forecasting tool, because of CART's use of "surrogate splits." The PM forecasting tools are based on the "primary splits" at the nodes of the decision trees created by CART. However, the CART tree also stores information on surrogate splits, which are rules for classification that are applied if the meteorological or air quality

variable used at the primary split is missing. In a real-time forecasting context, there are no missing variables because the forecaster can fill in datapoints with predictions or estimates. For the historical period evaluation described here, missing datapoints were not filled in and so CART resorted to surrogate data when necessary. This should be kept in mind when assessing the results of this subsection. Although the historical period evaluation may not use the exact same predictions the tools would have yielded in a real-time application, the predictions are probably similar. The results presented here are also of interest because, unlike the real-time 2004 evaluation, the summer 2003 period provides ample USG days to test the tools' ability to accurately predict high PM; furthermore, this assessment covers all nine MARAMA areas.

Table 5-3 provides several metrics for both the strict and fuzzy-border evaluations. Since the summer 2003 days were better distributed over Good, Moderate, and USG, the percentage of Good days is not given in the table as it does not provide the most useful comparison in this case. Because there are USG days for most areas, the accuracy, false alarm rate, probability of detection, critical success index, and bias are all informative measures of the tools' predictive utility. The table also provides the number of days evaluated for each area, as well as the number of strictly USG days.

Table 5-3. Evaluation Metrics for CART Historical Period Evaluation

The metrics are: Accuracy (Acc), FAR (False Alarm Rate), DetP (Probability of Detection), CSI (Critical Success Index), and Bias.

MARAMA Area	Days / USG Days	Strict Evaluation					Fuzzy-border Evaluation				
		Acc	FAR	DetP	CSI	Bias	Acc	FAR	DetP	CSI	Bias
Charlotte, NC	90 / 2	59%	na*	0%	0.00	0.00	72%	na*	na*	na*	na*
Bristol, VA	32 / 1	47%	100%	0%	0.00	1.00	53%	100%	0%	0.00	1.00
Roanoke, VA	31 / 1	52%	100%	0%	0.00	1.00	65%	100%	na*	na*	na*
Richmond, VA	89 / 2	71%	na*	0%	0.00	0.00	85%	na*	na*	na*	na*
Washington, DC	92 / 5	62%	88%	40%	0.11	3.40	71%	82%	50%	0.18	2.83
Baltimore, MD	92 / 8	57%	75%	25%	0.17	1.00	75%	50%	80%	0.80	1.60
Philadelphia, PA	85 / 4	66%	57%	75%	0.60	1.75	74%	57%	100%	0.75	2.33
Wilmington, DE	82 / 1	65%	88%	100%	0.14	8.00	85%	75%	100%	0.33	4.00
Newark, NJ	70 / 3	56%	80%	67%	0.22	3.33	63%	80%	100%	0.25	5.00

**Measure cannot be applied since no USG days were predicted and/or observed*

Predictive accuracy ranges from 47 to 71 percent if strict PM classifications are used, and 53 to 85 percent in the best-case scenario where borderline observations do not count against the tool. All areas except Charlotte have a bias (ratio of predicted to observed USG days) greater than one, and thus tend to overprediction, a fact also evident in the high false alarm rates. On the other hand, the probability of detection is fair to good for Newark, Philadelphia, and Wilmington, but problematic for the other areas. If only very high USG days are considered (and borderline USG days are counted as Moderate, according to the border zone definitions given

earlier in this section), the probability of detection is good or inapplicable for most sites, but still a problem for Bristol and Washington.

5.2.3. Conclusions

For a first attempt at developing a CART-based forecasting tool for these nine areas—the results are promising. The evaluation statistics are lower than but not that much lower than those that would be considered good for 8-hour ozone forecasting (and ozone is a simpler and much more extensively measured/studied pollutant).

The real-time testing of the draft version of the forecasting tools was inconclusive primarily because the period February-March 2004 did not contain any days with high $PM_{2.5}$ concentrations.

The historical evaluation suggests that given perfect forecasts of the meteorological input parameters, the $PM_{2.5}$ concentration ranges can be correctly predicted for 50 to 70 percent of the days and nearly corrected predicted (using the “fuzzy-border” adjustment) for 65 to 85 percent of the days (with the exception of Bristol, which has a 55 percent accuracy even with the adjustment).

In this historical evaluation, two of the sites with the worst performance are Bristol and Roanoke and these both had fewer data (with an every three day collection interval) than the other sites. Yet the CART trees for two sites had some of the best overall classification accuracy. This outcome suggests that the CART results, while good for characterizing the days in the dataset, are incomplete with respect to representing all of the types of conditions that might occur at these sites. The implication is that use of a limited dataset may limit the predictive ability of the tools, if conditions that are not represented in the dataset occur. This could extend to all areas and the use of the nominal three- to four-year analysis period.

The false alarm rate was relatively high for all areas, where it could be calculated, and this reflects the tendency for overestimation found in the CART trees. With this tendency, the probability of detection is good for most sites, and the bias is positive in all cases for which it could be calculated. This outcome suggests that the meteorological inputs and consequently the CART results may not sufficiently represent the conditions associate with the day-to-day transition from high to lower PM concentrations. The overpopulation of the higher PM bins with lower PM days (both in the CART results and in the historical forecast results) may also be due to a lack of a sufficient number of high PM days in the dataset, as needed to allow a good sampling and representation of the conditions that are associated with these days.

5.3. Operational and Research Versions of the Tools

The evaluation results, per se, did not lead to major revision of the tools. However, practice in using the tools, further consideration of the input parameters, and a few case studies by the state forecasters resulted in a few additions to the inputs. This further development of the tools is discussed in Section 2 of this report (CART diagnostic and sensitivity analysis) and resulted in a revised operational version of the tools as well as a research version of the tools that includes an estimated $PM_{2.5}$ concentration for the day prior to the forecast day.

This page intentionally left blank.

6. Summary and Recommendations

In this study, we developed a series of CART-based $PM_{2.5}$ forecasting tools for nine areas of interest in the MARAMA region including: Charlotte, Bristol, Roanoke, Richmond, Washington, Baltimore, Philadelphia, Wilmington, and Newark. The study included the application of CART and the development, testing, and evaluation of interactive forecasting tools for each area. Data and information gathered throughout the course of the project were used, together with the CART analysis results, to describe the relationships between meteorology and PM concentration and, specifically, the conditions associated with high $PM_{2.5}$ events in each of the areas. Based on the results and findings of the study, as well as the issues and problems that we encountered in conducting the work, we provide recommendations for future enhancement of the forecasting tools and an improved understanding of PM2.5 issues in this section.

The following recommendations pertain to the application of CART for $PM_{2.5}$:

- Update the input datasets to include additional years/seasons in order to better capture the range of different meteorological/$PM_{2.5}$ conditions that are likely to occur in the future as well as to better characterize the conditions associated with the high PM days (which were few in number during the analysis period for several of the areas).
- Using the expanded dataset with more high PM days, conclusively explore the use of alternative prior-day $PM_{2.5}$ concentration parameters for local and upwind sites, using both two-days-ago measured concentrations and prior-day estimated concentrations. It is intuitive that more information about the prior-day PM concentrations should improve the forecasting ability of CART, but our current work found the use of this information problematic (and resulted in the overestimation of $PM_{2.5}$ concentrations).

Additional recommendations pertain to the CART-based forecasting tools:

- Consistently (across the areas of interest) evaluate the forecasting tools for a longer period of time than was accommodated by this study. With a longer evaluation period, we may be able to identify specific patterns or types of PM events that are consistently missed by the CART-based forecasting tools. Combining and inter-comparing the evaluation results for the various areas of interest will aid the identification of missing parameters or information that is needed to capture the types of events that are consistently missed.
- Use the forecast evaluation results to reassess the uncertainty ranges used in the forecasting tools. These account for uncertainties in the input data (especially the meteorological forecasts) and their potential effects on the forecast.
- Evaluate and compare the use of the different meteorological forecast products (for example, ETA, GFS, and NGM).
- Add the capability for multi-day forecasts.
- Conduct detailed case-study analyses for as many of the high PM days as possible and compare the meteorologist perspective on important processes and parameters for the event with those used by CART to classify each day (i.e. generate the forecast)

Additional recommendations concern the improved understanding of the factors influencing $PM_{2.5}$ concentrations within each area of interest:

- Intermittently update the data summaries to include additional years/seasons of data.

- Examine, using available STN data, variations in species distributions among the CART bins and/or other groupings of the high PM days. This would need to be done using a larger dataset than that used for the current study – due to the more limited availability of STN data for the areas of interest.

Our final recommendations address the possible use of the data and results of this study to enhance $PM_{2.5}$ State Implementation Plan (SIP) analysis. For starters, the data analysis results for this study provide the basis for developing a conceptual description of $PM_{2.5}$ formation and transport for each area, which is a required component of a SIP. In addition, a key element of a $PM_{2.5}$ attainment demonstration is the "weight-of-evidence" analysis, in which data and modeling results are use to support or corroborate the outcome of the demonstration. The data analysis and CART results could be used to support the following types of weight-of-evidence analyses:

- Characterization of actual or proposed modeling episode periods in terms of their ability to represent typical meteorological conditions for each of the areas of interest. This would be determined based on the analysis of factors influencing PM2.5 in each area and the CART-based frequency of occurrence of the different types of meteorological conditions. This information could be used to guide the selection of an appropriate simulation period for the application of regional-scale particulate models, the selection of subset modeling episode periods for detailed analysis of certain areas, and the application of the modeled attainment test for $PM_{2.5}$.
- Analysis of data-based and meteorologically adjusted trends (adjusted using CART-based meteorological frequency information). Meteorologically adjusted trends, coupled with information about changes in emissions during the analysis period, could be used to assess the reasonableness of modeling results (i.e., the response of the model to similar emissions changes) and to project future changes in PM concentrations for the region.
- Calculation of meteorologically adjusted PM2.5 design values for use in the application of the $PM_{2.5}$ attainment test for each area of interest. Information from the CART analysis could be used to define a typical year (based on the frequency of occurrence of certain types of meteorological conditions) and the $PM_{2.5}$ design values corresponding to a typical multi-year period (based on actual observations).
- Use available Speciated Trends Network (STN) data in conjunction with the CART results to determine the species compositions of the most frequently occurring types of high PM2.5 events. This could help guide the identification of effective control options for the areas of interest or, in the context of weight-of-evidence, the interpretation and use of any modeling results.

References

AER, (2001) “Scoping Study for Regional Haze in the Upper Midwest,” Prepared for the Lake Michigan Air Directors Consortium, Des Plaines, IL, (Document No. CP089-01-1), February 2001.

Brieman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. *Classification and Regression Trees.* Wadsworth, Belmont, California.

Deuel, H. P., and S. G. Douglas. 1998. “Episode Selection for the Integrated Analysis of Ozone, Visibility, and Acid Deposition for the Southern Appalachian Mountains.” Systems Applications International, Inc., San Rafael, California (SYSAPP-98/07r1).

Douglas, S. G., J. B. Mangahas, A. B. Hudischewskyj, G. L. Glass, and W. S. Hartley. 2003. “Examination of Inter-Variable Relationships and Implications for Modeling Episode Selection Using the 1999/2000 SEARCH Database,” Systems Applications International, Inc., San Rafael, California (03-060).

EPA. (1998) “Atmospheric Observations: Helping Build the Scientific Basis for Decisions Related to Airborne Particulate Matter,” Report of the PM Measurements Research Workshop, Chapel Hill, North Carolina, 22–23 July 1998.

EPA. 2000. “Guidance for Demonstrating Attainment of Air Quality Goals for $PM_{2.5}$ and Regional Haze.” Draft Version 1.1.

Gillespie, W. G. 2004. “Correlating Federal Reference Method and Continuous $PM_{2.5}$ Monitors in the MARAMA Region.“ Final Draft Report, MARAMA, Baltimore, Maryland (040131).

Malm, W. C., (1999) *Introduction to Visibility* Cooperative Institute for Research in the Atmosphere (Cooperative Agreement CA2350-97-001), NPS Visibility Program, Colorado State University, Fort Collins, CO, May 1999.

McMillan, N. 2004. *Evaluation of $PM_{2.5}$ Surrogate Methods,* Battelle, Presented at the EPA 2004 National Air Quality Conference, Baltimore, Maryland, 22–25 February.

Seinfeld, J.H. (1986) *Atmospheric Chemistry and Physics of Air Pollution* John Wiley & Sons.

Steinberg, Dan and Phillip Colla. “CART—Classification and /regression Trees.” San Diego, CA: Salford Systems, 1997.

Watson J. G., (2002) *2002 Critical Review—Visibility: Science and Regulation,* J. of Air and Waste Management Association, Volume 52, 626-713.

This page intentionally left blank.

Conceptual Model of $PM_{2.5}$ Concentrations in Maryland

William F. Ryan
The Pennsylvania State University
Department of Meteorology
University Park, PA
wfr1@psu.edu

May, 2007

Abstract

Concentrations of $PM_{2.5}$ in the State of Maryland, while highest at urban monitors, are remarkably homogeneous across the state – particularly during the summer months. Maximum concentrations occur in the summer although urban monitors also observe a secondary, winter season peak. At all locations, annual mean concentrations are in excess of median concentrations due to a small number (~5-10%) of extremely polluted days. This subset of "dirty" days occurs primarily during the warm season (May-September) and is associated with light winds, strong low level inversions, regional scale transport of pollutants from west to east – similar to high O_3 episodes – and enhanced concentrations of sulfate particles. A similar transport pattern occurs in winter season cases also there is a higher frequency of stagnation in these cases as well as stronger near-surface inversions. Winter episodes thus feature a stronger influence of local emissions – especially carbon and nitrogen particles. Recently implemented regional control strategies, to the extent that they reduce sulfur and nitrogen emissions, may be effective at reducing $PM_{2.5}$ concentrations on the worst days although there remains significant local emission inputs to $PM_{2.5}$ in Maryland.

Executive Summary

1. This report summarizes observations of fine particulate matter with an aerodynamic radius of ≤ 2.5 μm, also known as $PM_{2.5}$, in the State of Maryland during the period 2000-2005.
2. Average $PM_{2.5}$ concentrations for this period at monitors across Maryland range from 12-17 μgm^{-3} with the highest concentrations observed at urban scale sites.
3. Although urban monitors observe the highest concentrations, all Maryland $PM_{2.5}$ monitors are strongly correlated and the correlation between monitors is highest during the summer season. This indicates that $PM_{2.5}$ is ubiquitous in Maryland, is not an overwhelmingly urban pollutant, and that all locations share in a common, regional scale "load" of $PM_{2.5}$.
4. Although there is no fool-proof method to quantify the regional scale "load" of $PM_{2.5}$, comparisons of urban, suburban and remote rural monitors suggest that the regional component of $PM_{2.5}$ accounts for roughly 60-75% of the total observed $PM_{2.5}$. This fraction increases to 80-90% during the summer season.
5. $PM_{2.5}$ concentrations peak during the summer season (June-August) in Maryland although urban scale sites also have a secondary maximum during the winter (December-February) months. The summer maximum are driven primarily by increases in the amount of sulfate while winter season peaks are driven more by increases in nitrogen and carbon compounds.
6. $PM_{2.5}$ concentrations vary by the day of the week, on the order of 2-3 μgm^{-3}, with highest concentrations occurring near the end of the work week and lowest concentrations on Sunday. This reflects day of week differences in motor vehicle and industrial emissions. While average concentrations do not vary significantly by the day of the week, the frequency of high $PM_{2.5}$ concentrations days (90^{th} percentile) is much greater during the work week.
7. $PM_{2.5}$ concentrations have a daily (diurnal) cycle with highest concentrations during the morning and afternoon rush hours, when emissions are highest and vertical mixing is weakest, and lowest concentrations during the well-mixed (diluted) afternoon hours.
8. The diurnal cycle is markedly different for the most severe (90^{th} percentile) cases. In those cases, the mid-day dilution effects are less evident so that concentrations remain nearly unchanged through the daylight hours. This effect is more pronounced in the summer months and suggests that the air aloft, which mixes downward in the afternoon, is heavily laden with transported $PM_{2.5}$.
9. The highest $PM_{2.5}$ cases are characterized overwhelming by westerly transport of air parcels although, in winter, there is a secondary maximum of cases where re-circulation, or stagnation, occurs. Observation at rural monitors west of Maryland show that, on the worst $PM_{2.5}$ days, this air mass is primarily made up of sulfate particles.

Introduction

This report summarizes observations of fine particulate matter with an aerodynamic radius of ≤ 2.5 μm, also known as $PM_{2.5}$, in the State of Maryland. $PM_{2.5}$ is regulated as a criteria pollutant by the US EPA with revised health and safety standards promulgated in 2006. The purpose of this report is to place the observations of $PM_{2.5}$ in the context of climate and weather conditions in order to aid policy makers in determining the best implementation plan to reach attainment with the $PM_{2.5}$ National Ambient Air Quality Standards (NAAQS).

Data

$PM_{2.5}$ has been measured routinely by a statewide network of monitors operated by the Maryland Department of the Environment (MDE) beginning in 1999. Due to data collection and quality issues associated with bringing a number of new monitors online during 1999, this study looks only at data collected during the period 2000-2005.

The majority of the data discussed in this report are from the network of Federal Reference Monitors (FRM) deployed across Maryland. The FRM $PM_{2.5}$ monitors are gravimetric, or passive, monitors that measure the integrated amount of $PM_{2.5}$ collected during a 24-hour period. Most monitors in Maryland collect data every third day and a subset collect daily. Statistical issues related to the uneven collection schedule were discussed in a previous report to MDE-ARMA. This report is included as ***Appendix A***. For the purposes of this report, daily monitor data are used whenever possible.

In addition to FRM monitors, MDE also operates continuous monitors that use a tapered element oscillating microbalance (TEOM) to measure $PM_{2.5}$ on short time scales – typically a one hour average. Continuous monitors are not currently designated as a FRM but, due to their good time resolution, provide useful information relevant to this study. Observations from the Old Town TEOM in Baltimore City are used at various places in this report.

Finally, because fine particles are made up of a mix of pollutants, MDE operates speciation monitors as part of the EPA Speciated Trends Network (STN). These monitors measure the individual components of $PM_{2.5}$, particularly sulfates, nitrates and carbon species. Observations from the Essex monitor, northeast of Baltimore City, are used in this report. In addition, a consortium of governmental organizations operates a network of speciation monitors as part of the Interagency Monitoring of Protected Visual Environments (IMPROVE) network. Monitors located in rural WV and Washington DC are germane to this report and data from these monitors are also used.

The location of the Maryland $PM_{2.5}$ network is shown in *Figure 1*. The Maryland monitor network includes sites that can be characterized as urban, suburban and rural. Some statistical measures within this report will analyze similarities and differences between observations at urban, suburban and rural sites and ***Appendix B*** provides a list

characterizing the location of the various Maryland $PM_{2.5}$ monitors. A related document (Power Point) provides location information and aerial views of many of the Maryland monitors (***Appendix C***).

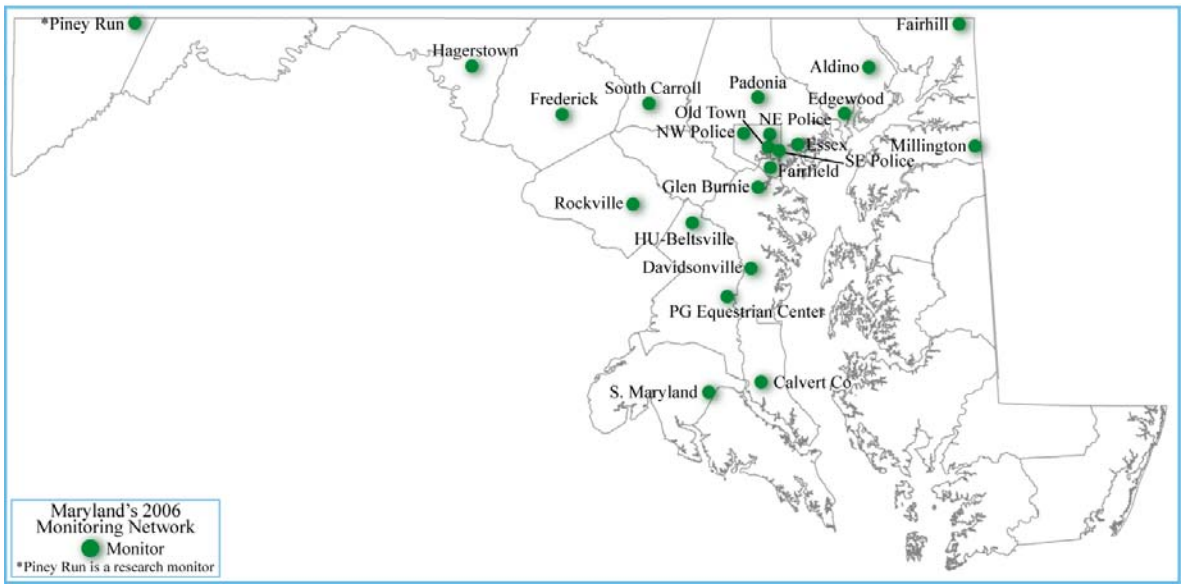


Figure 1. Maryland $PM_{2.5}$ FRM monitor locations.

In addition to the Maryland monitors, data from regional scale FRM $PM_{2.5}$ monitors are also referred to in this report. These monitors, located in PA, WV and VA, are designated as regional scale because they are typically located at high elevation in non-urban locations. The high elevation and non-urban locations of these monitors places them at a distance from concentrated emissions sources and provides a measure of $PM_{2.5}$ concentrations that are consistent with observations over a larger scale than an urban scale monitor can provide. Details on these monitors are provided in ***Appendix B***.

$PM_{2.5}$ Concentrations in Maryland (2000-2005 Average)

Mean $PM_{2.5}$ concentrations at the Maryland monitors for the period 2000-2005 are given in *Figure 2*. Monitors that re-located during the period, and do not have a complete, uninterrupted set of data, are not included. Average concentrations range from 12-17 μgm^{-3}. These concentrations are similar in magnitude to many eastern U.S. metropolitan areas and straddle the current yearly NAAQS for $PM_{2.5}$ (15.5 μgm^{-3}). As expected, the highest concentrations are found at the urban and near-urban monitors. As noted in ***Appendix A***, most monitors report data on every third day. The exceptions are three Baltimore urban monitors - Old Town, FMC and Essex - that report daily. The lowest concentrations are found in south suburban Maryland (Rockville and Davidsonville).

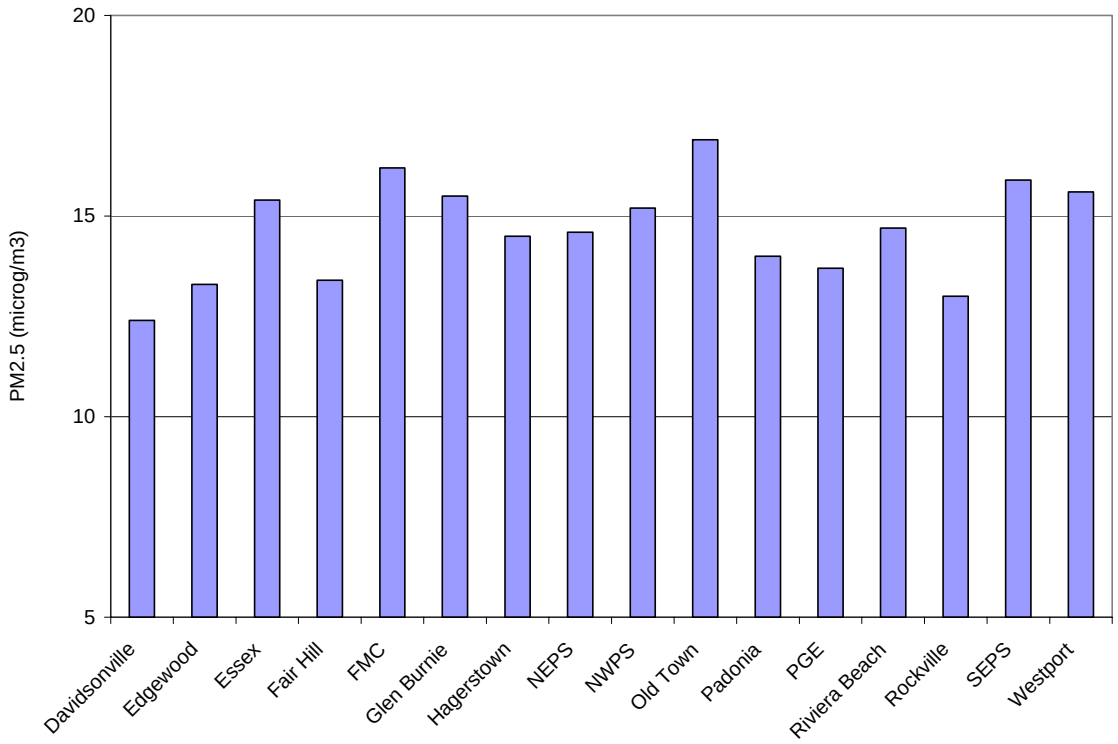


Figure 2. Mean $PM_{2.5}$ concentrations for the period 2000-2005 for Maryland FRM monitors. Certain monitors that are currently operational were re-located during this period and are not represented in this figure.

The urban sites, as defined in ***Appendix B***, have an average concentration of 15.7 μgm^{-3}, or slightly above the NAAQS. The suburban sites observed lower concentrations (13.6 μgm^{-3}). Hagerstown, a mid-size city in western Maryland, reports an average of 14.5 μgm^{-3} and the regional scale sites (not shown) report an average of 12.4 μgm^{-3}.

The most outstanding feature of the average $PM_{2.5}$ concentration in Maryland is the high degree of correlation between monitors across the state. *Figure 3* shows the correlation between all Maryland monitors and concentrations measured at the site with the highest average concentrations, Old Town – located near center city Baltimore. Even the far distant Hagerstown monitor shows a correlation coefficient of 0.78, and the remainder of the coefficients varies from 0.83-0.98. As an example, a scatter plot of $PM_{2.5}$ concentrations at Old Town and Fair Hill, Maryland – Fair Hill is located just south of the Pennsylvania border in extreme northeastern Maryland – is shown in *Figure 4*. For concentrations ≤ 25 μgm^{-3}, the agreement is extremely close but becomes less so at the higher end of the distribution.

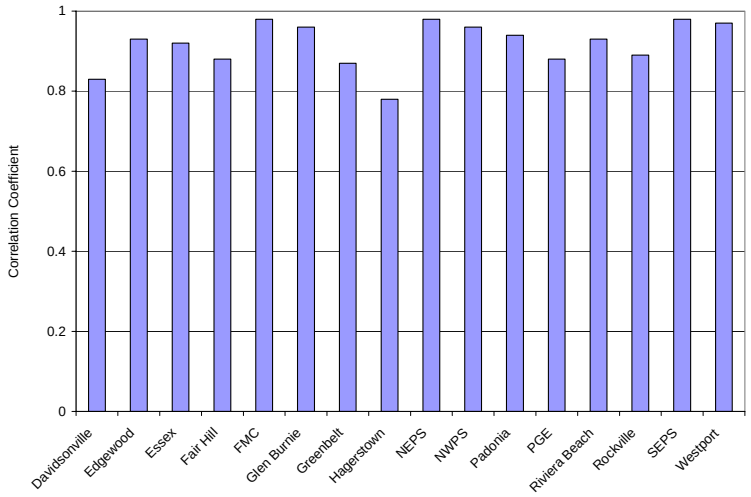


Figure 3. Correlation coefficients between 24-hour average $PM_{2.5}$ concentrations at Old Town, Maryland and the remainder of the Maryland FRM monitors.

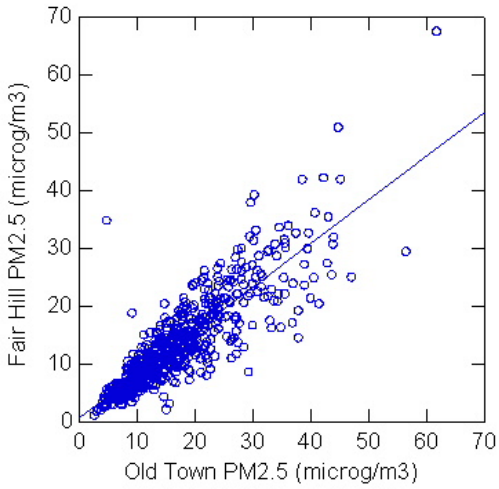


Figure 4. Scatter plot of 24-hour average $PM_{2.5}$ concentrations at Old Town, Maryland and Fair Hill, Maryland for the period 2000-2005. The best linear fit line is given as: $[PM_{2.5}]_{\text{Fair Hill}} = 0.78 + 0.75*[PM_{2.5}]_{\text{Old Town}}$.

The high correlation between all monitors suggests that there is some shared, regional scale "load" of $PM_{2.5}$. That is, a common source(s) of $PM_{2.5}$, either through primary emissions or secondary formation, that is ubiquitous and distributed evenly throughout the state. Primary emissions of $PM_{2.5}$, such as motor vehicle exhaust, tend to be short-lived with primarily local impacts and most strongly affect only nearby monitors. Secondary formation processes, such as the conversion of $SO_{2(g)}$ to particulate sulfate, occur on longer time scales and can affect larger areas downstream.

The primary component of $PM_{2.5}$ that accounts for the shared, regional load is sulfate. In *Figure 5*, the distribution of ammonium sulfate across the US is shown. Sulfate concentrations are highest, and similar in magnitude, across most of the eastern US and, as fraction of total $PM_{2.5}$, shown in *Figure 6*, account for approximately half of total $PM_{2.5}$ concentrations. The primary source of sulfates is coal combustion and these sources, typically electrical generation units, are spread throughout the region, not always in association with large cities. The distribution of SO_2 emissions, the precursors of sulfate, in the eastern US during the summer season are shown in *Figure 7*. The process of conversion from SO_2 as it leaves the stacks to sulfate particles occurs over many hours and the lifetime of sulfate is several days. Accordingly, sulfate can be transported long distances and attain homogeneous region-wide concentrations.

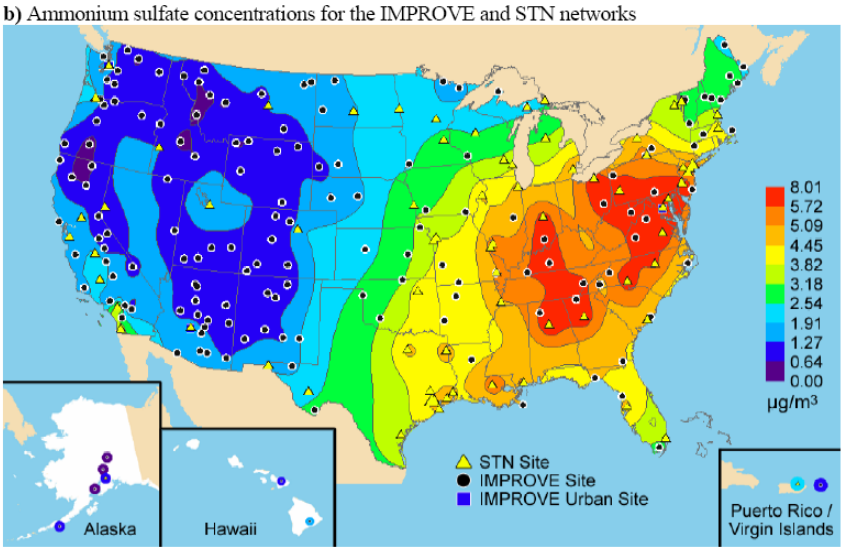


Figure 5. Ammonium sulfate concentrations measured at $PM_{2.5}$ speciation monitor sites ombining the rural IMPROVE monitors and the more urban scale STN monitors. http://vista.cira.colostate.edu/improve/Publications/Reports/2006/PDF/Chapter2.SpatialTrendsConcentration&Extinction.pdf

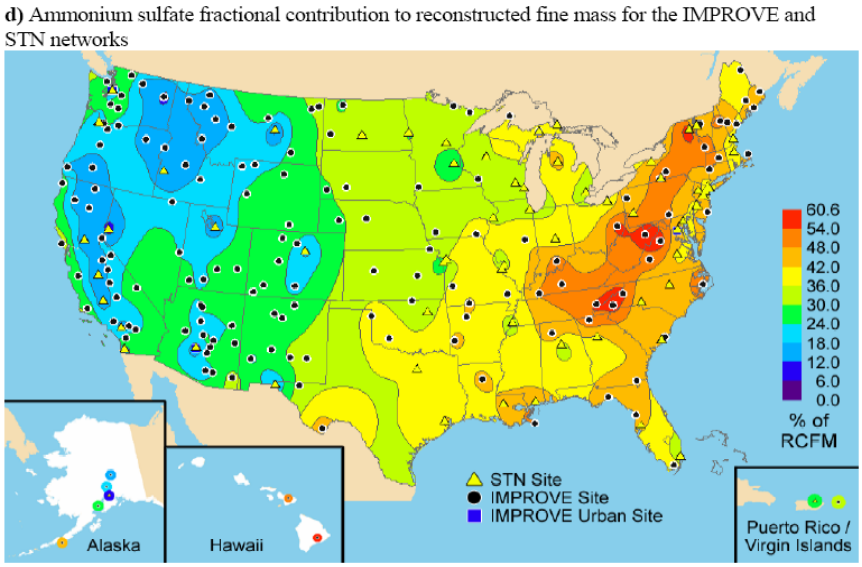


Figure 6. As in Figure 5 but showing ammonium sulfate concentrations as a fraction of total concentrations as measured at $PM_{2.5}$ speciation monitor networks (IMPROVE and STN).

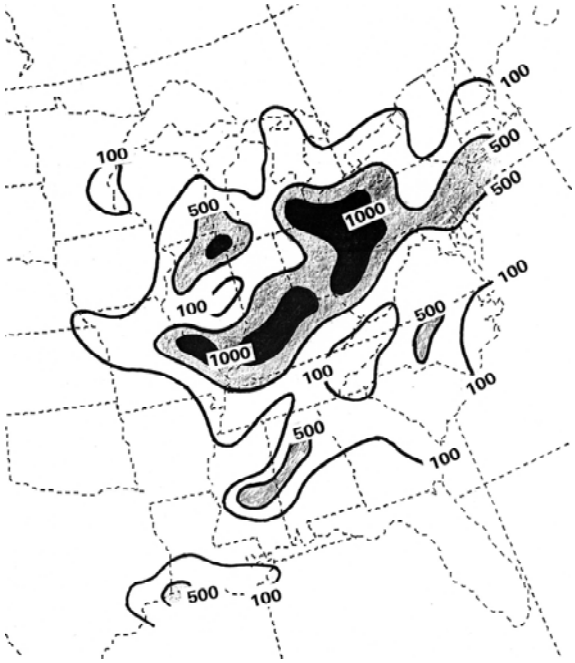


Figure 7. Distribution of SO_2 emissions during the summer over the eastern United States in metric tons per day (After Brueske, 1990). Source: http://www.spc.noaa.gov/publications/corfidi/fig3jpgmed.jpg

One way to quantify the magnitude of regional, or "background", $PM_{2.5}$, relative to local concentrations, is to determine the so-called "urban excess". That is, the amount of $PM_{2.5}$ that the urban environment adds to the regional load. The urban excess is often expressed as a simple paired comparison of remote and urban mean concentrations. An example from a recent EPA trends analysis is shown in *Figure 8*. Several eastern cities, including Baltimore, are shown on the right side of the figure. For this report, we pair concentrations at the remote Keeney Knob site in central WV with several local monitors (*Figure 9*). Compared to Old Town, the remote, regional scale concentrations represent 62% of the urban mean. At the suburban site at Padonia, north of Baltimore, the fraction increases to 75% and is 72% at Hagerstown.

In summary, average $PM_{2.5}$ concentrations in Maryland range from 12-17 μgm^{-3} with highest concentrations found in urban areas. $PM_{2.5}$ concentrations across the state are highly correlated, however, with a "regional scale load" of $PM_{2.5}$ that accounts for roughly 62-75% of the total observed $PM_{2.5}$ in urban areas.

Figure 12. Urban excess for total $PM_{2.5}$ gravimetric mass.

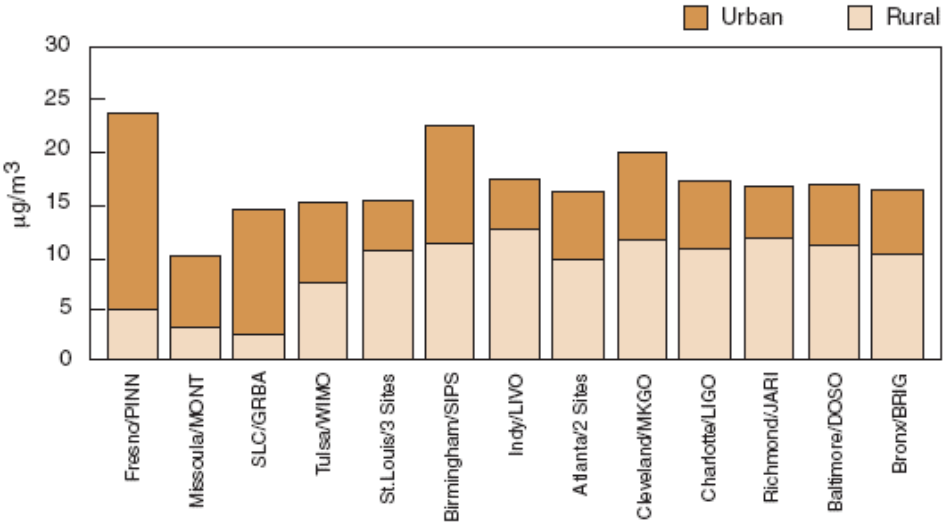


Figure 8. A comparison of urban and rural $PM_{2.5}$ concentrations as a meausure of "urban excess" $PM_{2.5}$. The second from the right column is Baltimore paired to Dolly Sods National Forest in remote WV. This figure and discussion is contained in http://www.epa.gov/air/airtrends/aqtrnd03/pdfs/2_chemspecofpm25.pdf.

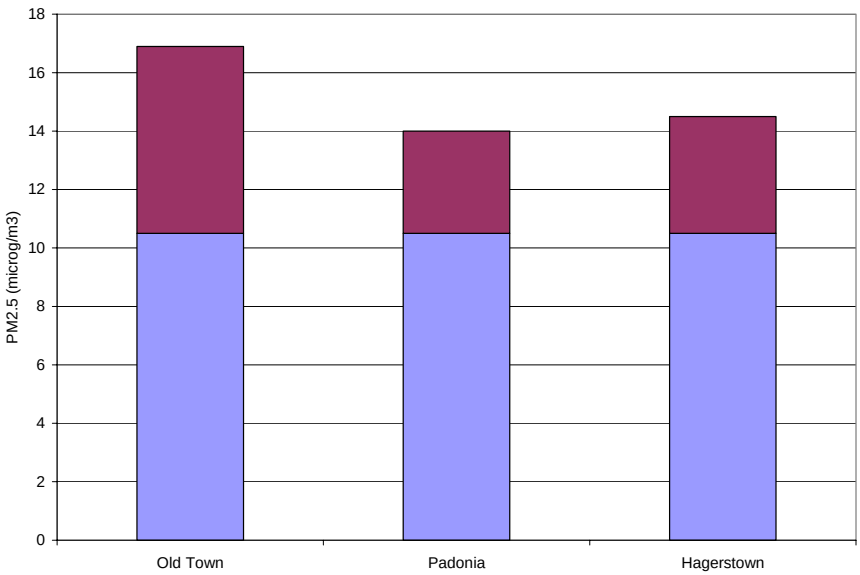


Figure 9. Mean concentrations at the remote, rural site located at Keeney Knob, WV (blue column) superimposed on mean concentrations at selected Maryland monitors. The difference, maroon column, is a rough measure of the "urban excess" or additional $PM_{2.5}$ added by local sources.

$PM_{2.5}$ Concentrations in Maryland (Seasonal Variations)

$PM_{2.5}$ concentrations in Maryland vary from urban to rural locations and also vary seasonally with highest concentrations occurring in the summer (June-August, JJA) months (*Figure 10*). Urban, suburban and central Maryland monitors, as well as more remote, rural sites in the region, all show peaks in the summer months. In addition, the Old Town (urban scale) monitor shows a secondary peak during the winter months (December-February, DJF). This seasonal pattern, and the differences between urban and rural concentrations, is consistent with prior measurements in the eastern US (*Figure 11*). At rural sites, like Shenandoah National Park and Dolly Sods National Forest, WV, the summer season maximum is pronounced with much lower concentrations observed for the remainder of the year. Urban monitors, such as Washington DC (*Figure 11*), and Old Town (*Figure 10*), have an additional, though smaller, winter season peak.

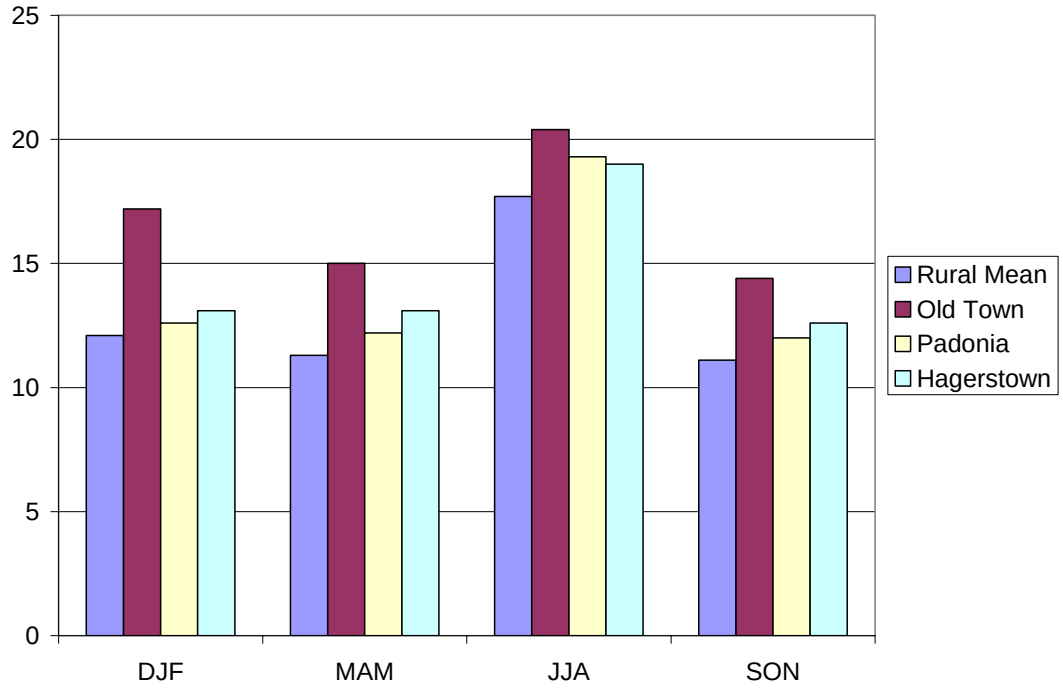


Figure 10. Seasonal average $PM_{2.5}$ concentrations at selected sites in Maryland (Old Town Baltimore, Padonia (suburban Baltimore) and Hagerstown (western MD)) for the period 2000-2005. In addition, seasonal average concentrations for a set of four monitors in VA, PA and WV ("Rural Mean") are given.

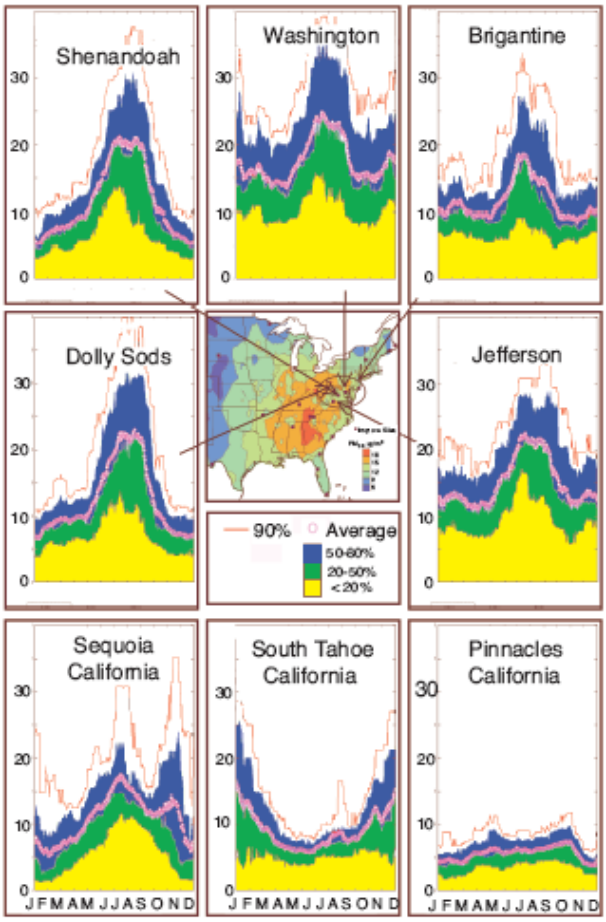


Figure 6.11. Seasonal variations of $PM_{2.5}$ mass concentrations at selected IMPROVE monitoring locations (1992-1999) for the cleanest (20 percent), moderate (20-50 percent), and highest PM (50-80 percent and 90 percent) days. The units of measurement are $\mu g/m^3$. (R. Husar, pers. comm.).

Figure 11. Seasonal concentrations of $PM_{2.5}$ at selected IMPROVE monitors for the period 1992-1999. Reprinted in:NARSTO (2004) Particulate Matter Assessment for Policy Makers: A NARSTO Assessment. P. McMurry, M. Shepherd, and J. Vickery, eds. Cambridge University Press, Cambridge, England.

At rural monitors, the summer season maximum in $PM_{2.5}$ is primarily driven by increases in ammonium sulfate ($[NH_4]_2SO_4$) concentrations. The efficiency of ammonium sulfate formation is maximized in a moist and highly oxidizing atmosphere, and so is more efficient in the summer months when humidity is high and sunshine plentiful. In *Figure 12*, speciated $PM_{2.5}$ concentrations are given for Dolly Sods Wilderness Area for 2004 (other years show similar patterns). The worst $PM_{2.5}$ days, labeled "W", all occur in the warm season and are all dominated by sulfate. In an urban environment (Washington, DC, *Figure 13*), summer season $PM_{2.5}$ is also dominated by sulfate but the winter season secondary maxima contain a significant fraction of nitrate and organic carbon. The larger concentrations of nitrate and organic carbon are not observed in rural environments due to much lower motor vehicle emissions.

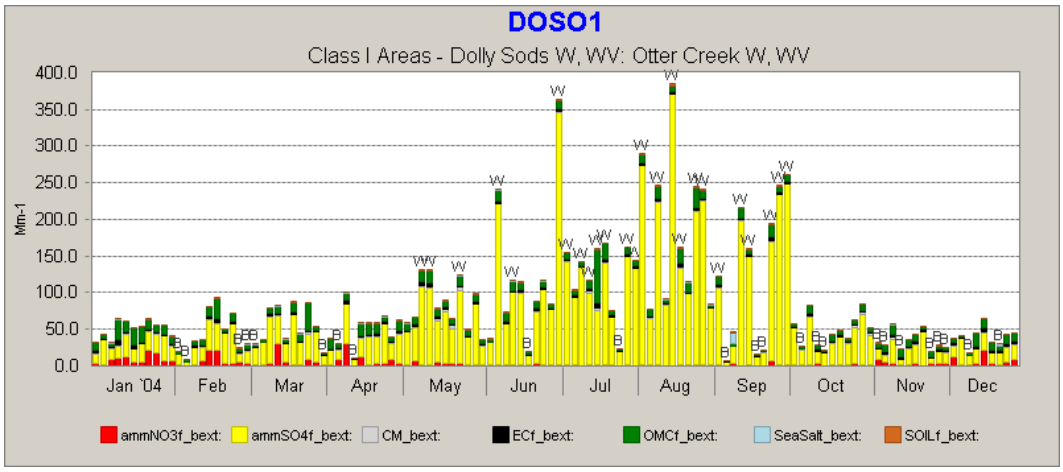


Figure 12. Speciated $PM_{2.5}$ concentrations for Dolly Sods Wilderness Area (WV) for 2004. Ammonium sulfate concentrations are in yellow. Data and plot courtesy of the Interagency Monitoring of Protected Visual Environments Program (IMPROVE) (http://vista.cira.colostate.edu/improve).

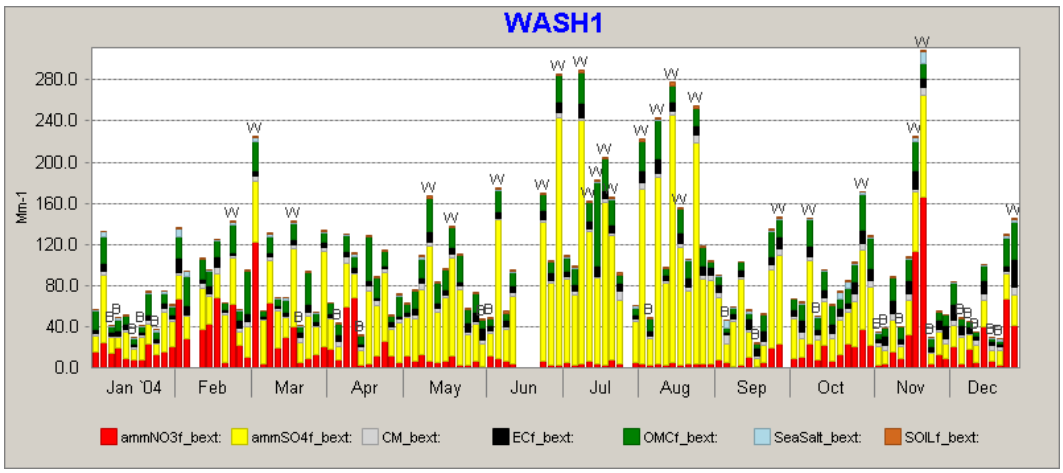


Figure 13. As in Figure 12, but for Washington, DC. Nitrate concentrations are given in red and organic carbon in green.

In Maryland, the effect of the summer season dominance of ammonium sulfate, as a fraction of total $PM_{2.5}$, is to make concentrations more homogeneous across the state. As noted above, sulfur emissions are primarily due to coal combustion from sources spread across the region with the strongest density to the west of Maryland and in the

Ohio River Valley (*Figure 7*). Because of the long lifetime of sulfate particles, they can travel long distances from their sources and become evenly spread across the region. As a result, correlation coefficients between urban, suburban and rural $PM_{2.5}$ in Maryland during the summer months, when sulfate is dominant, becomes greater (*Figure 14*, and compare to *Figure 3*). The "urban excess" of $PM_{2.5}$ in the summer months is correspondingly smaller. Comparing summer observations (*Figure 15*) with full year observations (*Figure 8*) we find that the regional load, which is roughly 62-75% overall, increases to 83-90% in the summer months. Returning to our example of Old Town and the distant exurban site at Fair Hill, we see a much tighter correspondence (compare *Figure 16* with *Figure 4*). The best fit line in summer has a slope of 0.91 compared to 0.75 for all days.

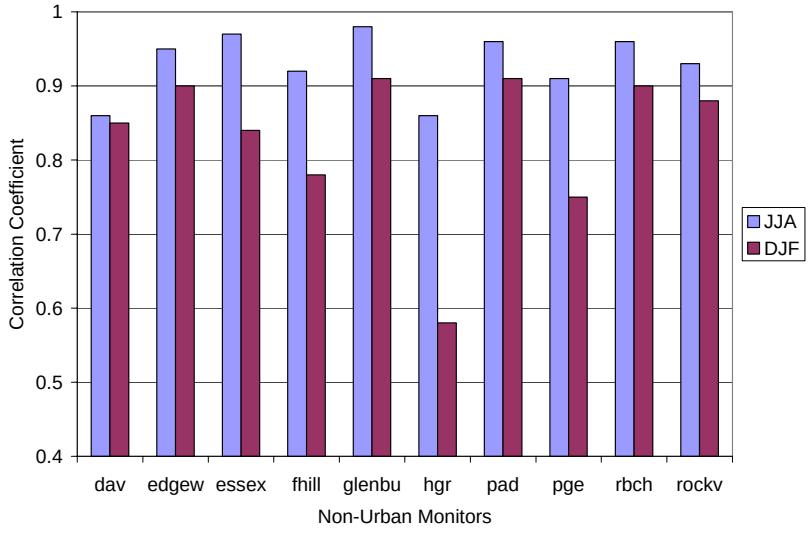


Figure 14. Correlations between $PM_{2.5}$ concentrations measured at Old Town (Baltimore) and non-urban locations in Maryland, as in *Figure 3*, but for only the summer (JJA) and winter (DJF) seasons.

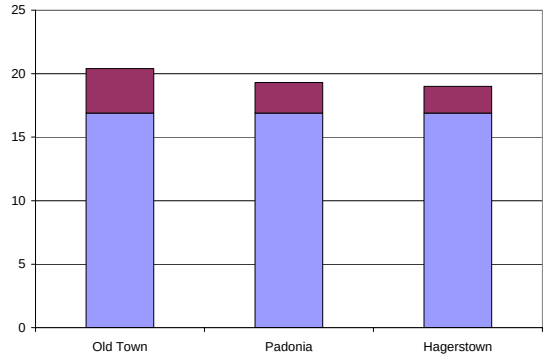


Figure 15. As in *Figure 8* but for only the summer season months (JJA). The ratio of Maryland monitors PM to Keeney Knob, WV increases by 34% for Old Town, 17% for Padonia and 24% for Hagerstown compared to the full year average.

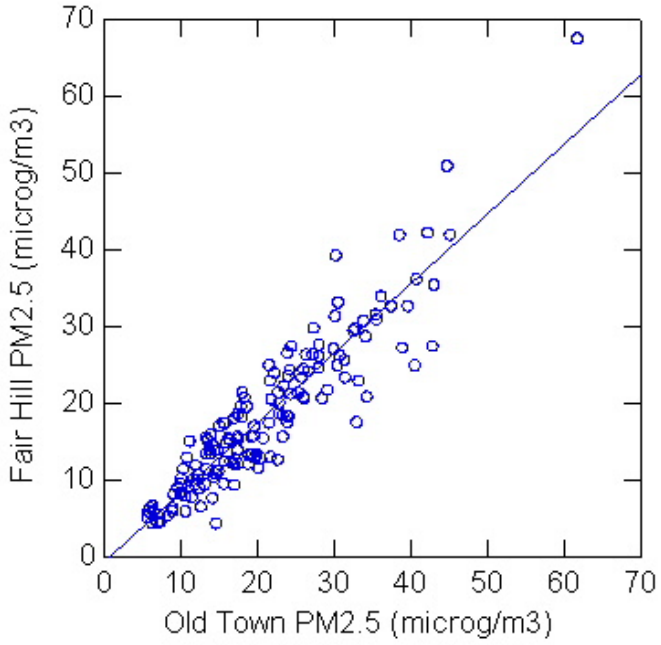


Figure 16. As in *Figure 4* but only for the summer season months (JJA). The best fit line is given by $[PM_{2.5}]_{Fair\ Hill} = [PM_{2.5}]_{Old\ Town}*0.91 - 0.59$.

In summary, $PM_{2.5}$ concentrations are highest during the summer season across Maryland. This summer maximum is characterized by a large, regional scale sulfate load with rural locations observing 83-90% of the urban scale concentrations. Urban scale monitors also observe a secondary winter season maximum that is not observed at rural locations. The winter season $PM_{2.5}$ is associated with significant concentrations of nitrate and carbon compounds.

$PM_{2.5}$ Concentrations in Maryland (Daily Variations)

$PM_{2.5}$ concentrations in Maryland vary seasonally and also by day of the week. The highest concentrations are typically observed on weekdays and peak in the Thursday-Friday period (*Figure 17*). This day of week difference is consistent with motor vehicle usage. As an example, in *Figure 18*, vehicle miles traveled (VMT) for a six week period in the Detroit area is shown. VMT is calculated as the number of vehicles using the system times the distance they travel. For the time period displayed in *Figure 18*, Sundays are the low points on the graph. On weekdays, the trend toward highway travel later in the week (Thursdays and Fridays) is common in most urban areas. While commuting trips are relatively stable throughout the week, discretionary trips are higher as the weekend approaches. The day of week plot for Maryland (*Figure 17*) suggests that day of week emissions variations can account for ~ 2 μgm^{-3} difference in $PM_{2.5}$ concentrations in Maryland. This difference is consistent across seasons.

Although higher weekday emissions lead to only slightly higher mean concentrations, day of the week differences appear to be important during the most severe $PM_{2.5}$ cases. In *Figure 19*, the day of week breakdown for the highest 90^{th} percentile ($PM_{2.5} > 30.75$ μgm^{-3}) of $PM_{2.5}$ days statewide is given. The weekday average number of days in the 90^{th} percentile is 47 while the weekend days are lower – Saturday is 34 and Sunday 22. This suggests that while weekday emissions differences have a small overall impact on $PM_{2.5}$ concentrations, they can be a significant influence during $PM_{2.5}$-conducive weather conditions.

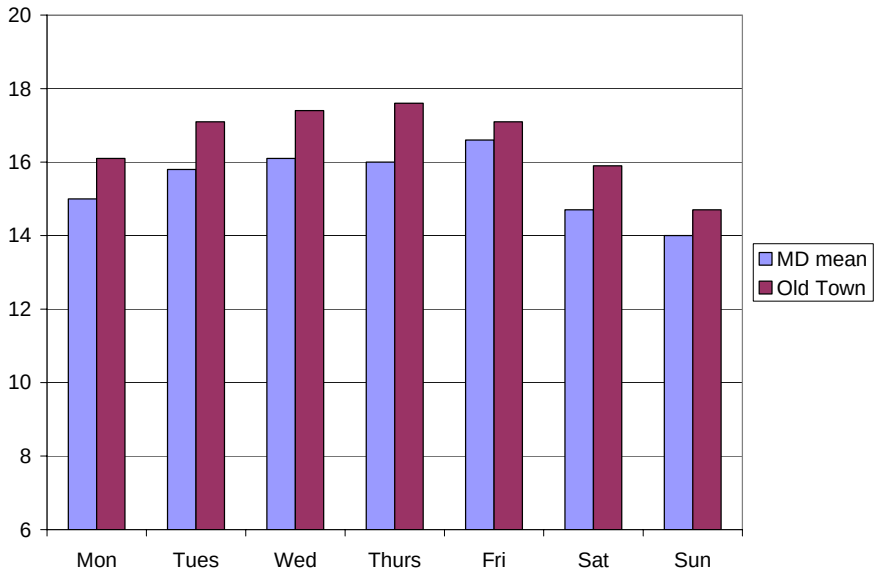


Figure 17. Day of week average $PM_{2.5}$ concentrations for statewide Maryland monitors and for Old Town.

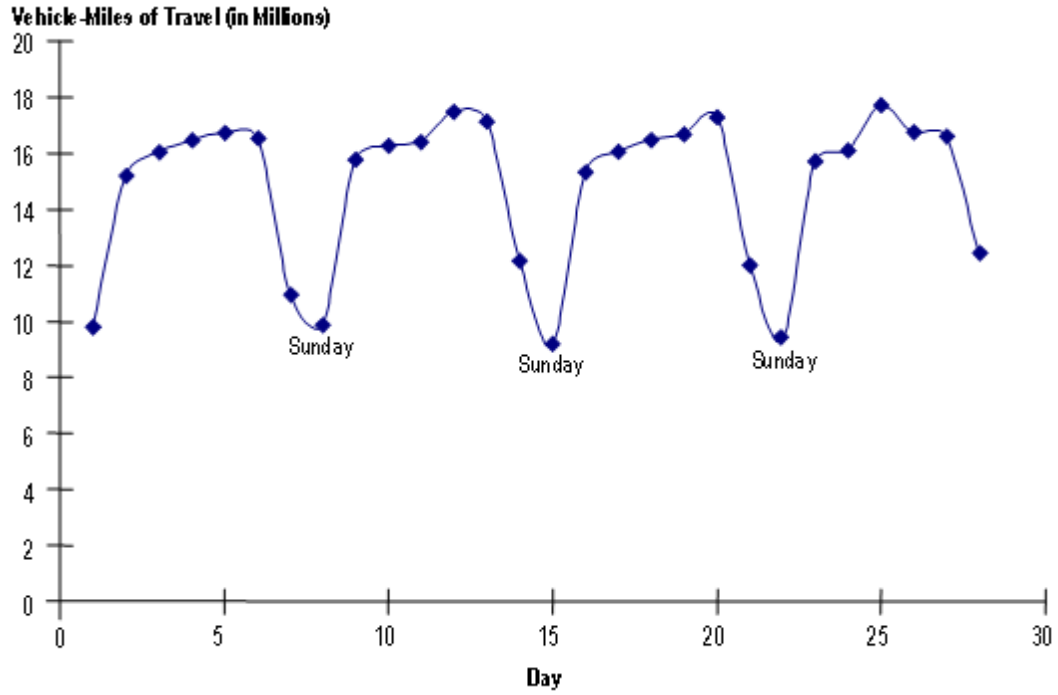


Figure 18. Graph of vehicle-miles of travel (VMT) for the Detroit metropolitan freeways for the period March 11-April 7, 2001. VMT is a common measure of highway usage (http://ops.fhwa.dot.gov/congestion_report/chapter2.htm).

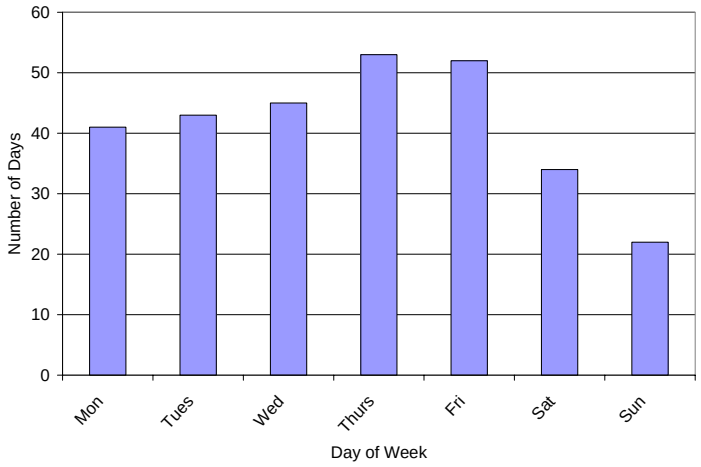


Figure 19. 90th percentile $PM_{2.5}$ days in Maryland by day of week.

$PM_{2.5}$ concentrations in Maryland (Intra-Day Variations)

Daily cycles of $PM_{2.5}$ concentrations in urban areas show the influence of local sources as well as the effects of transported pollutants. The diurnal pattern shows a two peak pattern that reflects emissions and mixing effects (*Figure 20*). The morning rush hour, when motor vehicle emissions are maximized, is shown as a sharp peak in concentrations during the early morning hours. In addition to increasing in strength as more cars are on the road, the emissions are also trapped within a shallow boundary layer. The boundary layer, or mixing depth of the atmosphere, is typically at a minimum during the early morning hours. This follows from surface-atmosphere differences in heat conduction. The ground is a good conductor of heat while the atmosphere is not. As a result, the ground cools quickly overnight while the atmosphere near the ground cools slowly. A surface-based inversion develops as a result of this gradient in temperature, with warm air over cool land surfaces, and traps pollutants within a layer just several hundred meters deep. By mid-day, heating of the earth's surface breaks the inversion and deep vertical mixing follows. This mixing tends to dilute the polluted air parcels with cleaner air from aloft and concentrations fall. This trend to lower $PM_{2.5}$ during the well-mixed afternoon hours is seen in *Figure 20*. Concentrations rise again toward evening through a combination of increased emissions during the afternoon rush hour and decreased mixing as the sun sets and the ground cools. As particles dry deposit to the surface during the nighttime hours, and emissions are low, concentrations fall again to a minimum just before the morning rush hour.

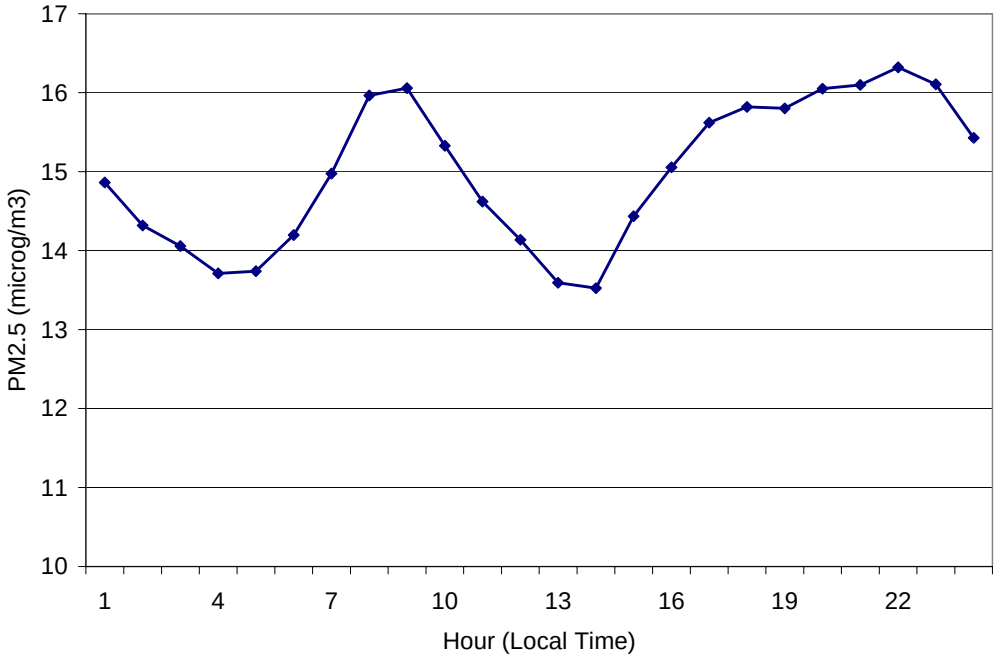


Figure 20. Time series of hourly $PM_{2.5}$ concentrations measured by a continuous monitor at Old Town, Maryland.

The diurnal cycle shown in *Figure 20* is changed subtly in high $PM_{2.5}$ cases. In *Figure 21*, hourly $PM_{2.5}$ concentrations during the 90th percentile cases are shown. The 90th percentile cases show a similar two-peak pattern but with important differences. First, the morning maximum extends until mid-day. This could result from either a stronger low level inversion, trapping pollutants in a narrow layer for more hours, or a dirty residual layer – the air resident above the nighttime boundary layer - mixing downward. The period of lower mid-day $PM_{2.5}$ during the well-mixed afternoon hours in the 90th percentile cases lasts only for 3 hours before concentrations return to near the morning levels while, for the 0-90th percentile cases, lower concentrations last for 8 hours from ~ 1000 to ~ 1800 local time. Another way of looking at the difference is given in *Figure 22*. In *Figure 22*, the ratio of hourly concentrations to the daily maximum is given. For the 0-90th percentile cases, the two rush hours, where motor vehicle emissions dominate, are clearly demarcated. In the high $PM_{2.5}$ cases, concentrations rise through the usually well-mixed afternoon hours.

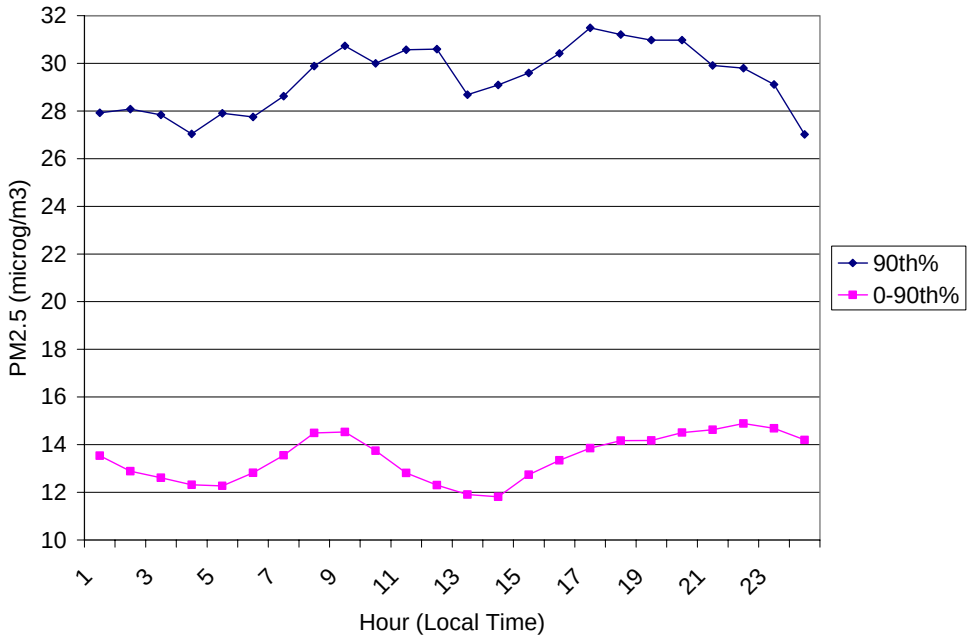


Figure 21. Hourly $PM_{2.5}$ concentrations at Old Town, Maryland for the 90th percentile and above cases (dark blue, diamonds) and the 0-90th percentile cases (magenta, square).

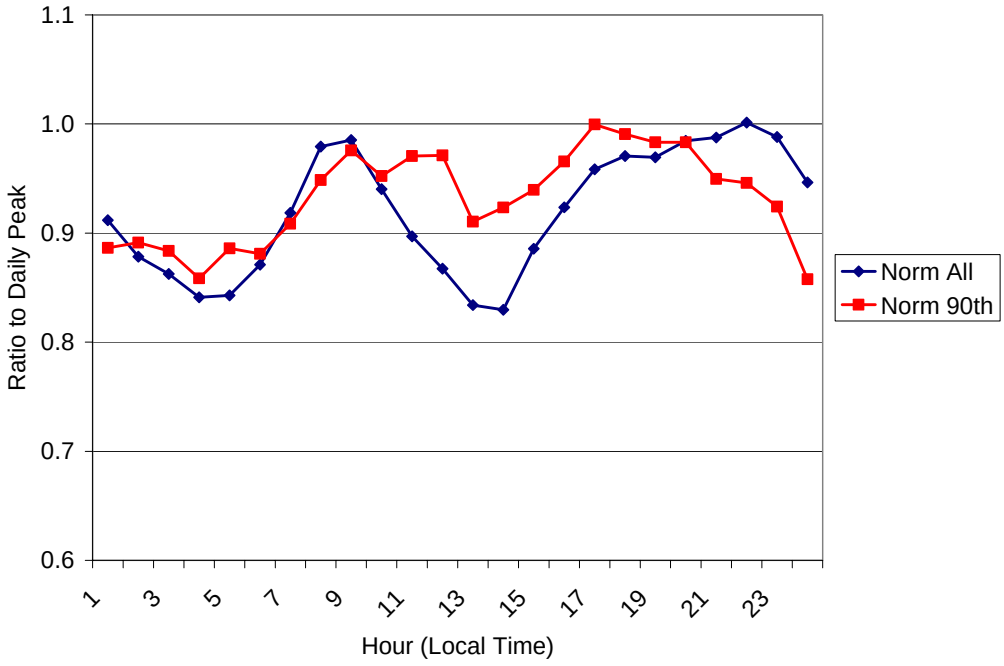


Figure 22. As in *Figure 21*, but with hourly concentrations normalized by ratio to the daily peak concentrations.

Disaggregating the data further provides more insight into the nature of high $PM_{2.5}$ cases in Maryland. In *Figure 23*, the 90th percentile cases are further divided into summer and winter season cases. In summer, the overall concentrations tend to be higher but, more importantly, the diurnal pattern is quite different. In the winter cases, the influence of the low level inversion is stronger. Note that concentrations remain high until mid-day. At that point, mixing reduces $PM_{2.5}$ levels with a modest rise during the evening rush hour. In summer, however, concentrations dip slight immediately after rush hour, but, as mixing proceeds in the afternoon hours, concentrations actually rise. This indicates that significant $PM_{2.5}$ concentrations are present aloft and are mixing downward. If that were not the case, dilution due to deeper mixing would reduce concentrations.

In summary, hourly observations at urban scale Maryland monitors show two daily peaks coincident with the morning and afternoon rush hours. As vertical mixing maximizes during the warmer mid-day hours, concentrations fall to a value consistent with average regional scale concentrations. During high $PM_{2.5}$ days, the daily pattern is significantly different and varies by season. During summer high $PM_{2.5}$ events, regional scale concentrations are very high – that is, there is little "urban excess" – so that $PM_{2.5}$ concentrations fall only slight, if at all, during the well-mixed afternoon hours. In winter, the regional loading is weaker but a stronger morning inversion leads to an extension of the morning rush hour effect.

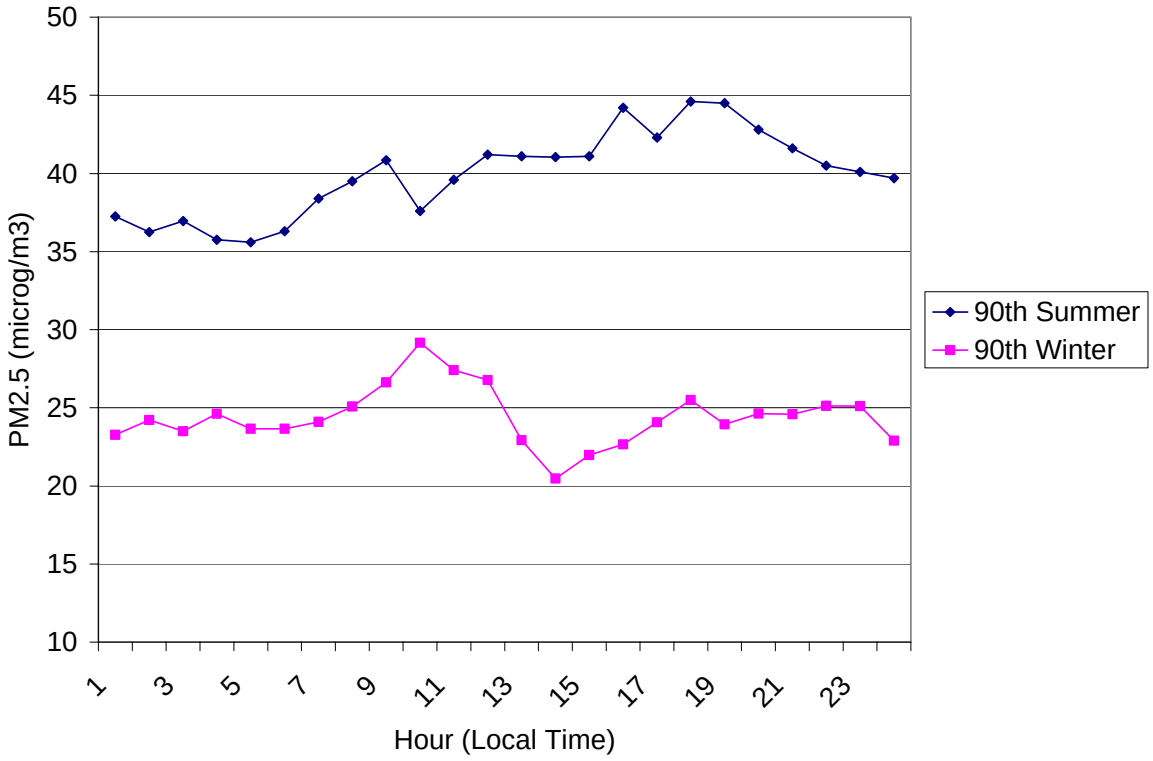


Figure 23. Hourly average $PM_{2.5}$ concentrations at Old Town for the 90th percentile summer and winter cases in Maryland.

$PM_{2.5}$ in Maryland: High $PM_{2.5}$ Cases

As noted above, $PM_{2.5}$ concentrations follow a different daily pattern in the highest (dirtiest) cases and, in particular, during the summer months. Are these cases somehow unique? One of the common attributes of $PM_{2.5}$ at all locations is that the distribution of concentrations is not normal.[1] An example for Old Town is given in *Figure 24*. The skewed nature of the distribution at Old Town is common to all Maryland $PM_{2.5}$ monitors. A distribution with a strong right-tail, as is the case here, results in the mean of the distribution being well in excess of the median. As *Figure 25* shows, this is the case for all Maryland monitors.

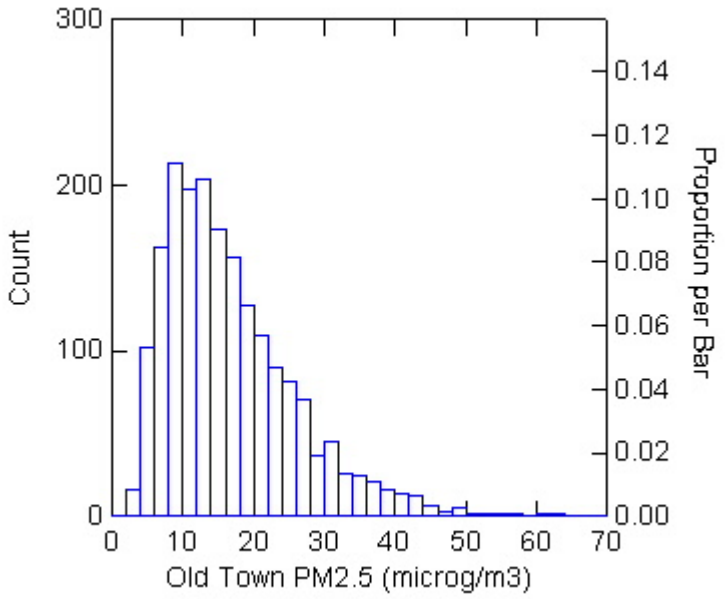


Figure 24. Daily average $PM_{2.5}$ concentrations at Old Town, Maryland.

[1] The distribution of $PM_{2.5}$ concentrations for Old Town, as well as for all monitors in Maryland, are significantly different from normal with longer than expected right tails (asymmetric). For Old Town, the skewness measure, a measure of the symmetry of a distribution about its mean, is significantly positive and nonzero (skewness = 1.19). A skewness coefficient is considered significant if the absolute value of skewness divided by the standard error of skewness (SQR(6/n)) is greater than 2 – for Old Town, this value is 19.8. In the same manner, Old Town, as well as all Maryland monitors have much larger extremes (tails) than a normal distribution. The measure of the length of the tails of the distribution is kurtosis. For Old Town, the measure of kurtosis is 1.67 with a standard error of kurtosis of 0.11. The absolute value of kurtosis divided by the standard error of kurtosis is 15.2 and as such is significantly greater than 2.

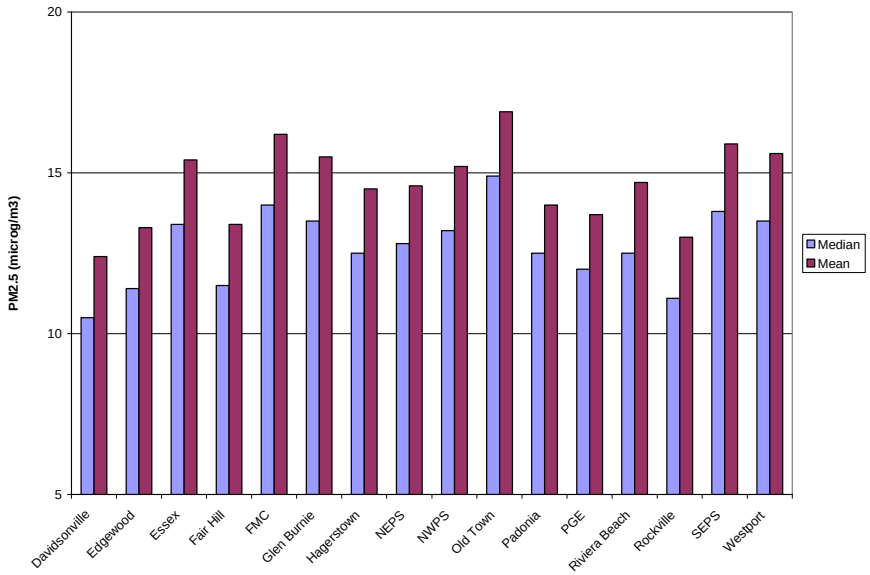


Figure 25. Median and mean $PM_{2.5}$ concentrations for Maryland monitors. The higher mean concentrations indicates that the overall distribution is right-skewed with a larger than normal subset of high $PM_{2.5}$ cases.

The implications of a highly skewed distribution are several. First, it suggests that there may be more than one physical process contributing significantly to the $PM_{2.5}$ distribution. That is, the high $PM_{2.5}$ cases may result from factors that are not as strongly present in the "normal" cases. This is to be expected in the eastern US where there can be large influences on any given day from any of the major constituents of $PM_{2.5}$. Sulfate can dominate, particularly in summer, while nitrogen and carbon, which reflect motor vehicle emissions, tend to dominate in the urban environment. Second, compliance with the NAAQS, based on annual mean concentrations, may depend on the presence or absence of the extreme cases. For example, mean concentrations at Old Town, as well as the other two daily reporting sites in Baltimore, are above the yearly standard of 15 μgm^{-3}. Removing the highest 10^{th} percentile of $PM_{2.5}$ days ($\geq$ 29.4 μgm^{-3}) reduces the yearly mean to 14.3 μgm^{-3}. There are also implications for daily $PM_{2.5}$ forecasting. The 95^{th} percentile of the 2000-2005 $PM_{2.5}$ distribution at Old Town is 35.2 μgm^{-3}. The other daily reporting sites in Maryland, Essex and FMC are similar (32.1 μgm^{-3} and 34.4 μgm^{-3} respectively). These concentrations are close to the newly proposed daily NAAQS for $PM_{2.5}$.

In the following section, we will look in detail at seasonal high $PM_{2.5}$ episodes but we can here make some general observations. In particular, the transport pattern associated with the worst $PM_{2.5}$ days is consistent. In *Figure 26*, back trajectories for Baltimore for the 95^{th} percentile of $PM_{2.5}$ days are shown. The back trajectories estimate the path of air parcels arriving in the Baltimore area. Overall, slow westerly transport is

the rule. There are a handful (9) of cases with winds from the east and a slightly larger number of re-circulation cases, but the vast majority of cases involve westerly transport. As noted in *Figure 7*, significant sources of SO_2 emissions are located west of Baltimore along with major large industrial centers.

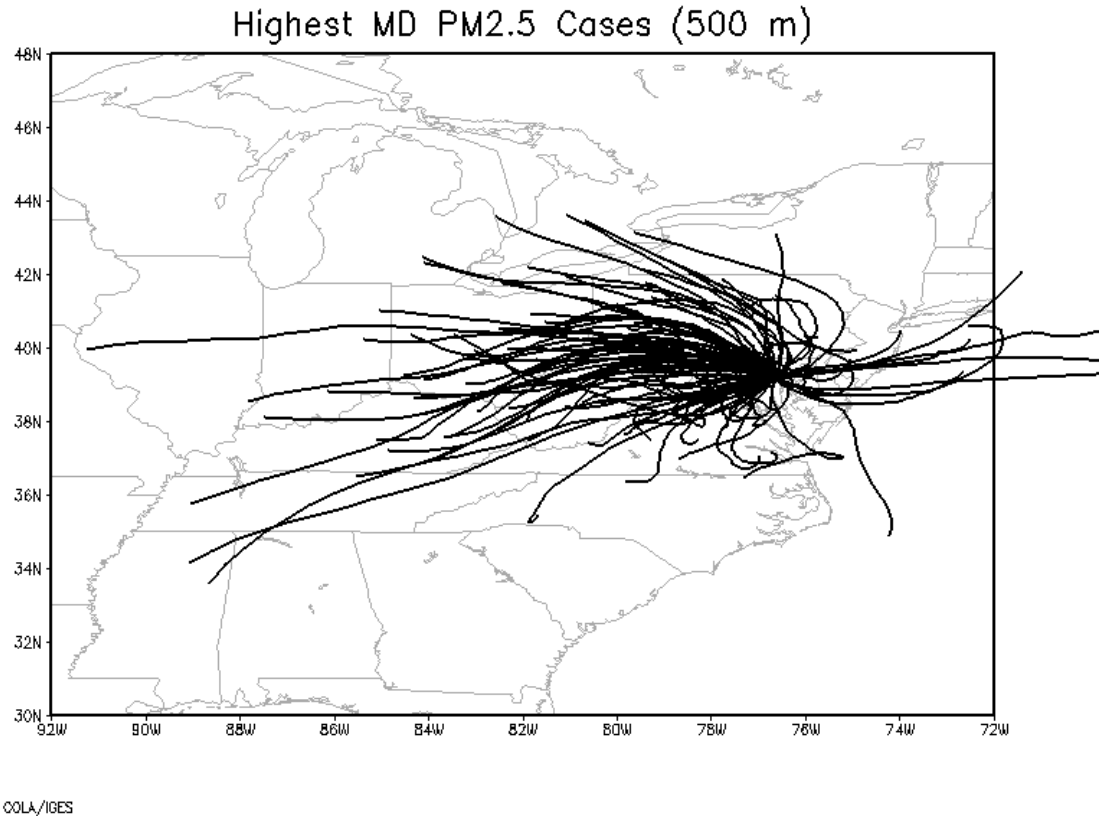


Figure 26. Back trajectories for the 95th percentile of pm cases in Maryland. The back trajectories are determined using the NOAA HYSPLIT model (http://www.arl.noaa.gov/ready/hysplit4.html) . The back trajectories shown here, and in succeeding figures terminate at BWI at 500 m above ground level (agl) at 1200 UTC (0700 EST) and are for 24 hours of travel time.

There are seasonal differences in transport patterns. In *Figure 27*, back trajectories for only warm season (May-September) high $PM_{2.5}$ cases are shown with winter cases shown in *Figure 28*. As a fraction of the total, the winter cases are less dominated by westerly transport with approximately one-third of the winter cases exhibiting some degree of re-circulation or stagnation of air masses. During the summer season, recirculation accounts for only ~ 15%. During the worst $PM_{2.5}$ cases, therefore, the transport pattern is overwhelming westerly. The make up of air parcels arriving from the west is primarily sulfate. As shown in *Figure 29*, concentrations at Dolly Sods, upwind of Maryland in westerly transport that characterizes poor $PM_{2.5}$ days, is primarily composed of sulfur compounds.

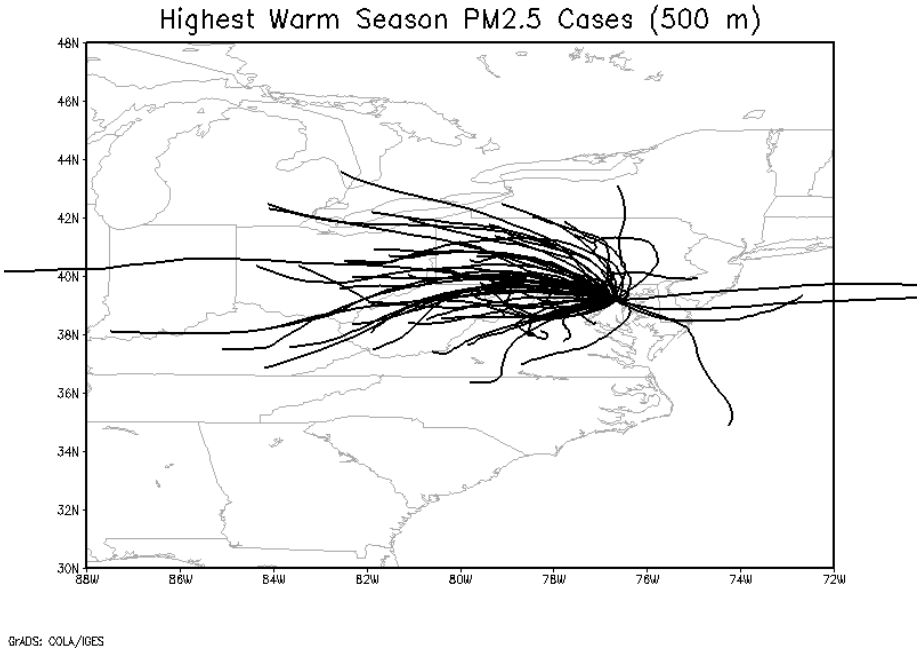


Figure 27. As in *Figure 26* but for warm season (May-September) cases.

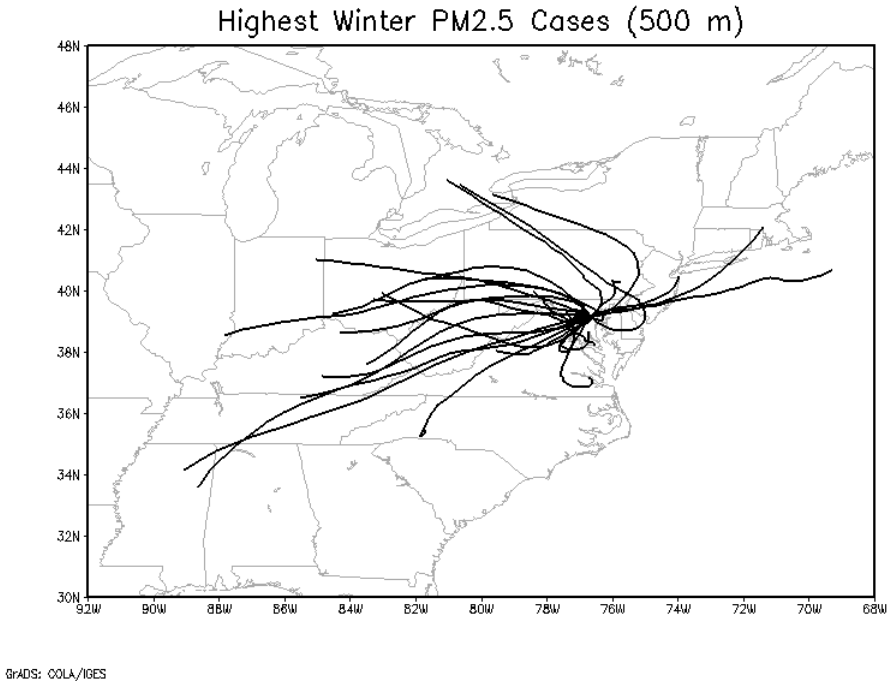


Figure 28. As in *Figure 27* but for winter (December-February) cases.

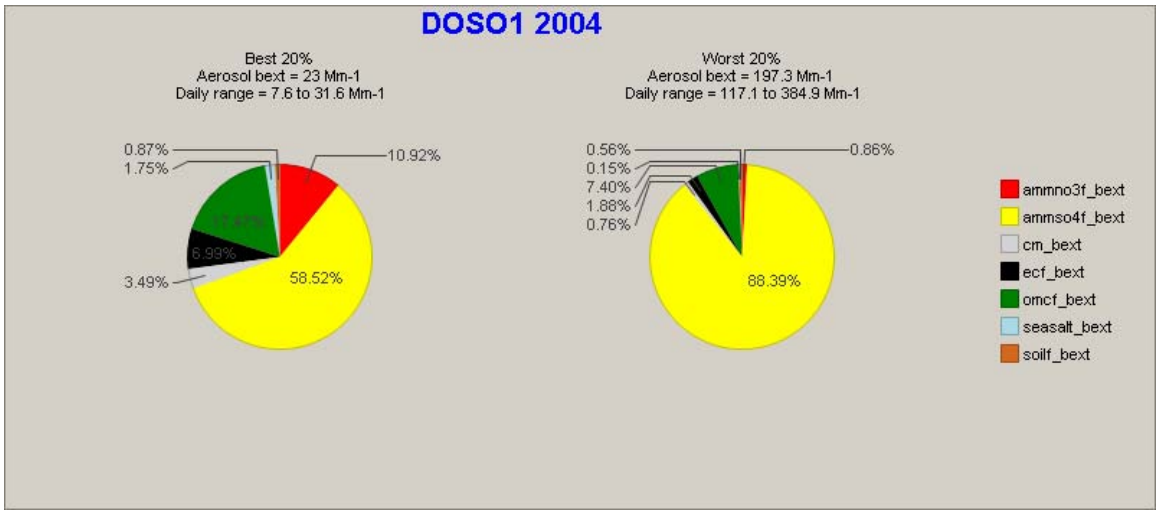


Figure 29. Speciated $PM_{2.5}$ measurements from the IMPROVE monitor at Dolly Sods, WV, for 2004 during the best (cleanest) days (left) and the worst (dirtiest) days (right).

In summary, the distribution of $PM_{2.5}$ concentrations in Maryland is right-skewed with mean concentrations strongly affected by a small fraction of very high $PM_{2.5}$ days. These days are predominantly characterized by westerly transport aloft, particularly during the summer season. Winter season high $PM_{2.5}$ cases show a higher incidence of stagnation along with westerly transport.

$PM_{2.5}$ Episodes: Meteorological Influences

Summer Season

$PM_{2.5}$ and O_3 concentrations in the summer season are highly correlated (*Figure 30*) and meteorological conditions associated with summer season multi-day $PM_{2.5}$ episodes are similar in nature to severe O_3 episodes. This coincidence follows for several reasons. First, $PM_{2.5}$ and O_3 share sources and precursors. For example, volatile organic compounds (VOCs) form a significant fraction of $PM_{2.5}$ mass and are also a critical O_3 precursor. NO_x, while a small fraction of $PM_{2.5}$ in summer, is generally the limiting factor for O_3 and a large percentage of regional scale NO_x concentrations have their source in coal-burning power generation units that also produce sulfur compounds - the main fraction of summer season $PM_{2.5}$. Second, weather conditions conducive to O_3 formation are also conducive to $PM_{2.5}$ accumulation. In particular, any weather processes that limit horizontal or vertical mixing will increase concentrations of both pollutants. Local (limited mixing) and regional (westerly transport) scale processes that increase $PM_{2.5}$ can both be maximized by weather patterns that feature light surface winds, strong, or multiple inversions, and westerly transport aloft.

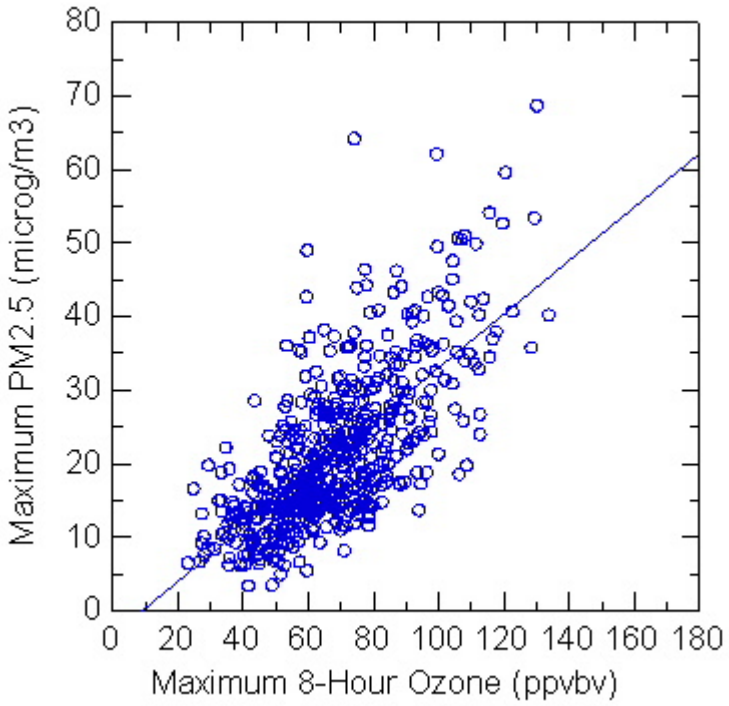


Figure 30. Peak daily 8-hour O_3 and 24-hour $PM_{2.5}$ concentrations in the Baltimore metropolitan area (Summer, 2000-2005).

O_3 and $PM_{2.5}$ conducive weather patterns do diverge in some important respects. O_3 formation is strongly dependent on incoming UV radiations to drive the photo-chemical reactions that produce O_3. Sunny skies mean more O_3. While sunlight is helpful for the oxidation of some $PM_{2.5}$ precursors, primarily SO_2 and organic carbon particles, it is a less important factor in relative terms. As a result, Maryland can observe high $PM_{2.5}$ concentrations, with low O_3, in the presence of significant cloud cover. In addition, while O_3 is not strongly correlated with relative humidity – high relative humidity typically means a better chance for cloud formation, $PM_{2.5}$ is associated with high humidity regimes. The conversion of $PM_{2.5}$ precursor gases, especially SO_2, to particles is enhanced in high moisture environments, and this process is enhanced further by cloud droplet processing.

Taken together, the large scale weather conditions conducive to summer season $PM_{2.5}$ formation are often associated with weather patterns also conducive to O_3, while smaller (meso) scale effects can vary the relative strength of O_3 and $PM_{2.5}$ on any given day. The large scale factors include: (1) An upper air ridge with its major axis west of Maryland. Subsidence, or downward motion, occurs downwind (east) of the ridge axis leading to decreased cloud formation and less vertical mixing; (2) With a ridge west of Maryland, winds aloft are westerly and tend to advect air masses high in $PM_{2.5}$ precursors from the Ohio River Valley and Midwest; (3) Surface high pressure typically leads the upper air ridge axis by a quarter of a wavelength so that the center of surface high pressure is overhead when the upper air ridge is west of Maryland. With high pressure overhead, temperatures increase, pressure gradients are weak and surface winds are light. This allows local pollutants to accumulate and mix with pollutants transported into the area aloft.

Summer $PM_{2.5}$ Episode: July 15-22, 2002

The interaction of weather with $PM_{2.5}$ in the summer season will be shown with reference to a strong $PM_{2.5}$ episode during July 2002. This episode is chosen because it met the criteria for severe $PM_{2.5}$ events (***Appendix D***) and enhanced $PM_{2.5}$ monitoring was carried out in Maryland during that month with the Essex site obtaining daily $PM_{2.5}$ speciation data. Each high $PM_{2.5}$ event is, of course, slightly different, but this event featured most of the key factors. Enhanced $PM_{2.5}$ levels were observed from July 15-22 (*Figure 31*). Peak concentrations exceed 50 μgm^{-3} on July 17-18 and, while enhanced on the other days of the episode, are in the 20-30 μgm^{-3} range. In addition, mean statewide and peak concentrations are quite close in magnitude, an indication of the regional scale of this event.

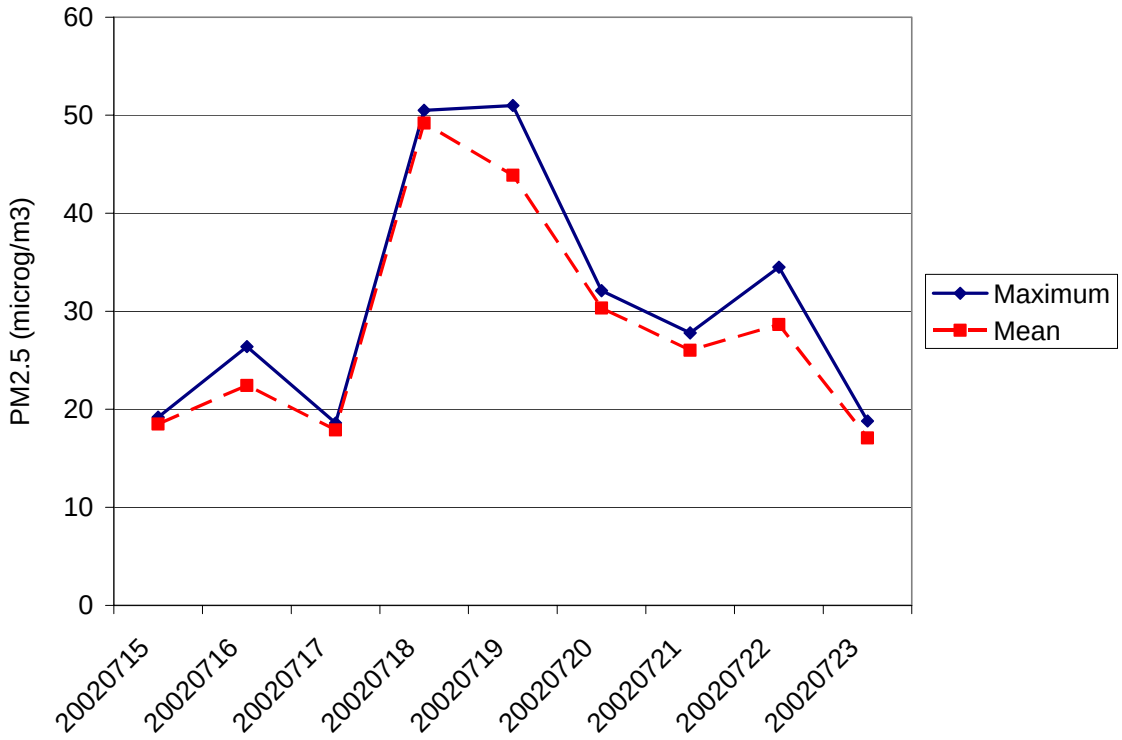


Figure 31. Daily $PM_{2.5}$ concentrations in Maryland during the high $PM_{2.5}$ episode of July, 2002. Concentrations from the monitors with the maximum $PM_{2.5}$ is given in blue and the average of all Maryland monitors in red.

The weather conditions associated with this episode are similar to the standard form with an upper air ridge of high pressure to the west (*Figure 32*), surface high pressure overhead (*Figure 33*) and generally west to northwest transport of pollutants (*Figure 34*). Ozone is enhanced throughout the period (*Figure 35*) as expected for this weather pattern. Note, however, that O_3 concentrations show little day-to-day variability when compared to $PM_{2.5}$.

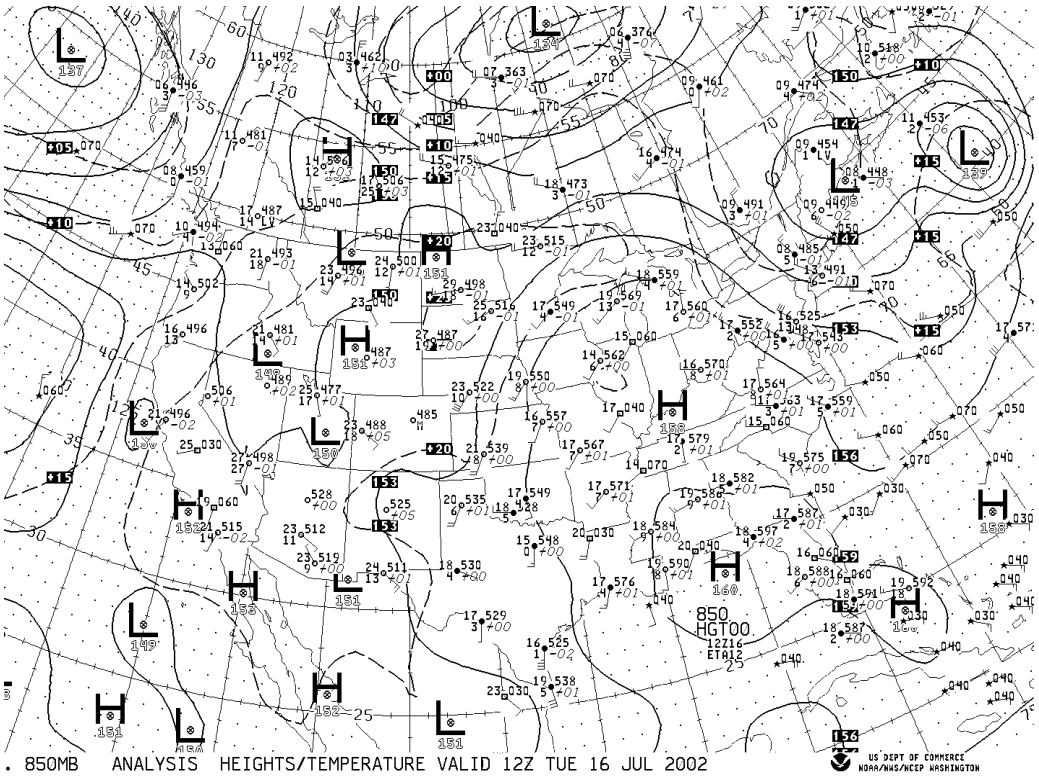


Figure 32. NCEP analysis of 850 mb weather variables for 1200 UTC on July 16, 2002. Contours are geopotential height, with station data showing height, temperature, dew point depression, height tendency and wind speed and direction.

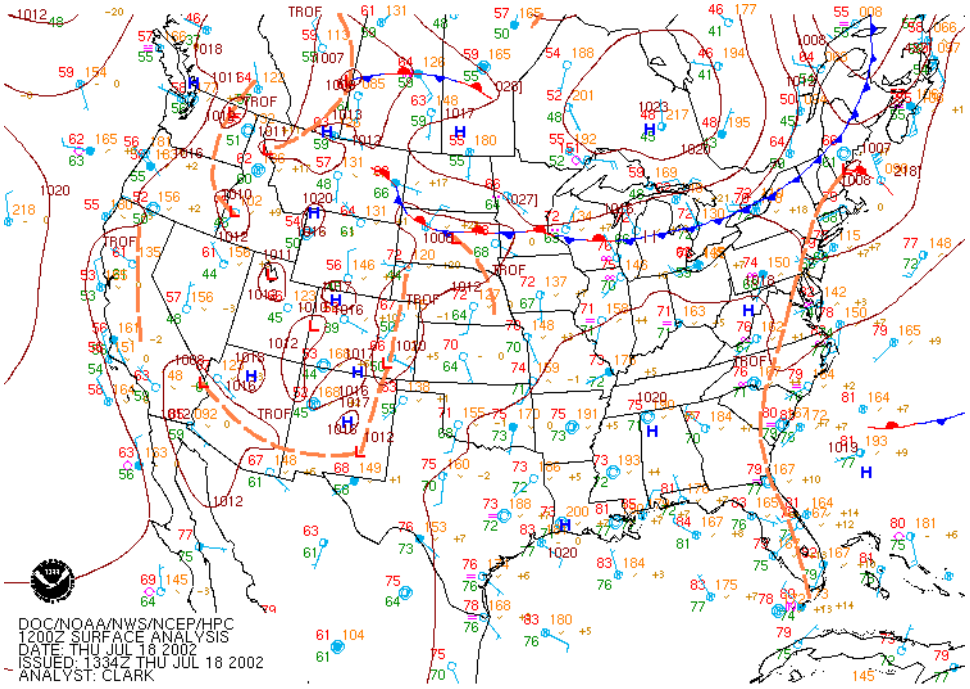


Figure 33. NCEP surface analysis for 1200 UTC on July 18. 2002.

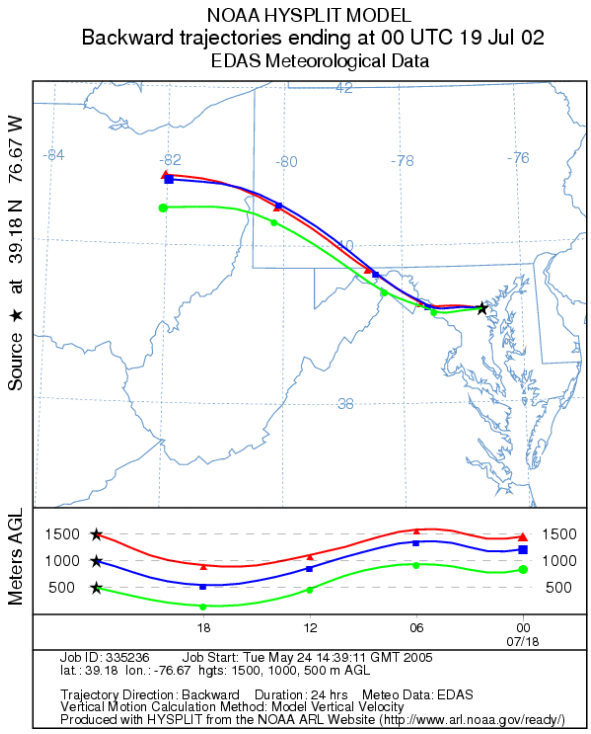


Figure 34. HYSPLIT back trajectories terminating at BWI for 0000 UTC on July 19, 2002. Trajectories are for 24-hours of travel and terminate at 500 (green line), 1000 (blue line) and 1500 (red line) meters agl.

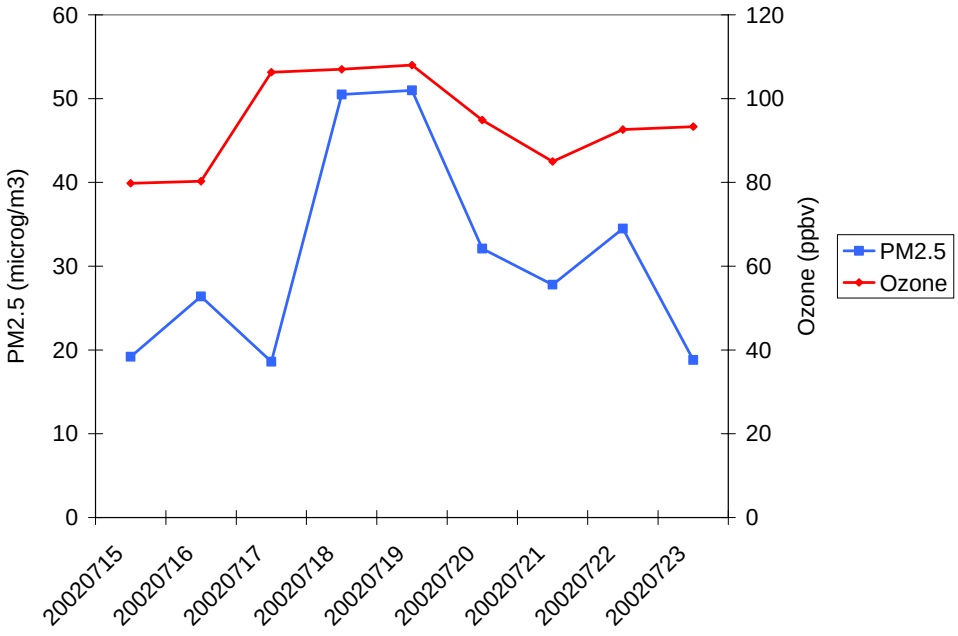


Figure 35. Peak O_3 (8-hour average) in Baltimore and $PM_{2.5}$ (24-hour average) in Maryland for July 15-23, 2002.

What accounts for the strong peak in $PM_{2.5}$ concentrations? From the Essex data, it is apparent that the peak in $PM_{2.5}$ concentrations is associated with increased sulfate levels. The key source of sulfate is coal combustion and it appears that the westerly transport during that period introduced significant sulfate levels to Maryland. Hourly observations on July 18-19 show that $PM_{2.5}$ concentrations tended to rise (*Figure 36*) during the well mixed afternoon hours indicative of high pm concentrations in the residual layer aloft. Daily speciated $PM_{2.5}$ observed at Essex, Maryland (*Figure 37*) show that nearly all of the observed increase in $PM_{2.5}$ during July 18-19 is due to increases in ammonium sulfate. Hourly observations on July 18-19 show that $PM_{2.5}$ concentrations rose during the well mixed afternoon hours (*Figure 36*) indicative of significant $PM_{2.5}$ within the residual layer aloft. Daily speciated $PM_{2.5}$ observed at Essex (*Figure 37*) shows that nearly all of the increase in pm on July 18-19 is due to increases in ammonium sulfate.

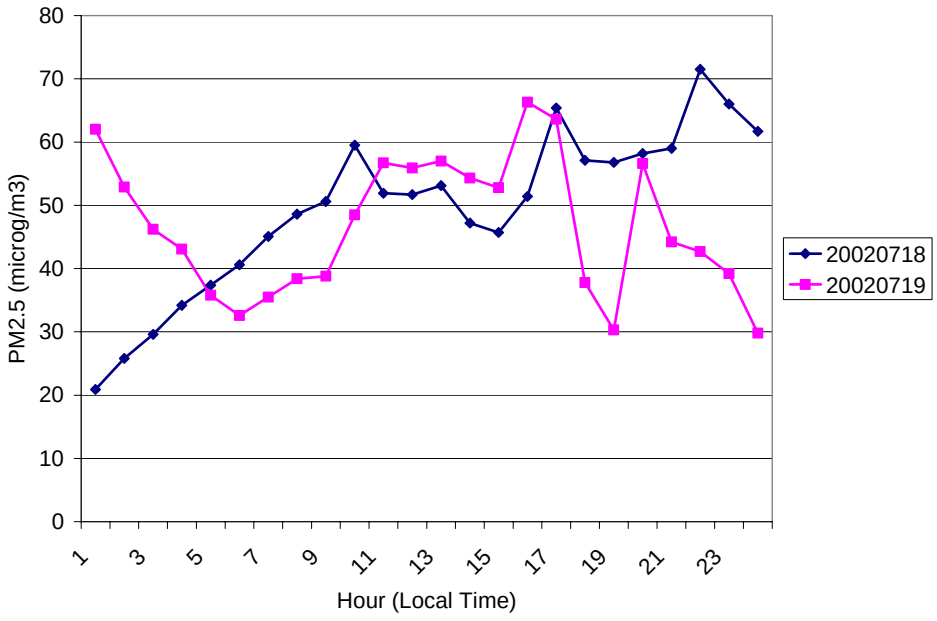


Figure 36. Hourly $PM_{2.5}$ concentrations at Old Town, Maryland for July 18-19,2002.

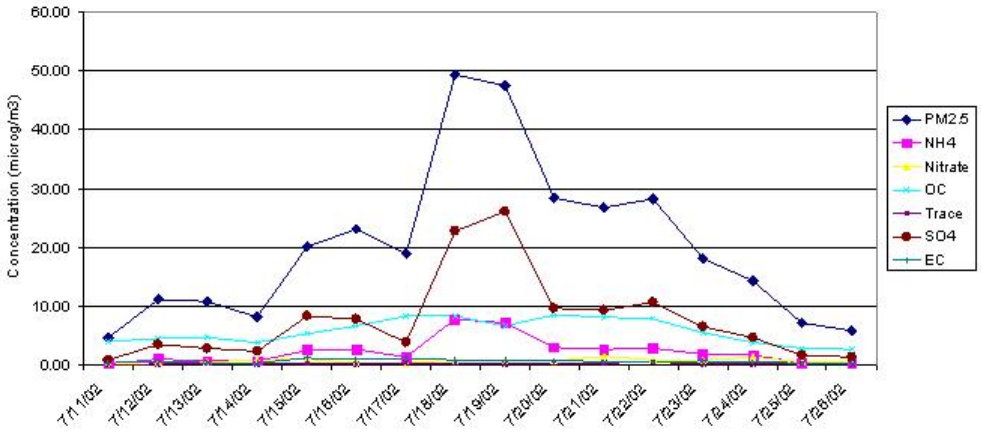


Figure 37. Daily average concentrations of speciated $PM_{2.5}$ observed at Essex, Maryland for July 11-28, 2002.

In summary, summer season $PM_{2.5}$ episodes feature weather patterns quite similar to high O_3 events. Westerly transport aloft provides a rich source of regional scale $PM_{2.5}$ and its precursors and light near surface winds allow local emission to accumulate as well. On the most extreme $PM_{2.5}$ days, evidence suggests that the bulk of the excess $PM_{2.5}$ is composed of sulfate compounds.

Winter Season $PM_{2.5}$ Episodes

High $PM_{2.5}$ days are less frequent in winter than summer (*Table 2*) and, while having many similar characteristics with respect to weather patterns, are also more complex with respect to $PM_{2.5}$ constituents. This difference may have implications for the efficacy of control strategies.

Frequency of Days with Maximum $PM_{2.5} \geq 35\ \mu gm^{-3}$		
Season	Number of Days	Percentage
DJF	42	29%
MAM	14	10%
JJA	66	46%
SON	22	15%
Total	144	100%

Table 2. Breakdown of high $PM_{2.5}$ days in Maryland by season for the period 2000-2005.

The most important difference in winter season cases is the greater likelihood of stagnation occurring during the pollution episode. In *Figure 38* and *Figure 39*, composite plots of average weather conditions during winter season high $PM_{2.5}$ cases during 2000-2003 are shown. The average sea level pressure field (*Figure 38*) shows a lobe of high pressure centered over the Atlantic coastal region. Weak pressure gradients are associated with high pressure overhead and this leads to light and variable near-surface winds. In *Figure 39*, very light winds near the surface (1000 mb) winds are associated with locations near the high pressure center. For the 90^{th} percentile $PM_{2.5}$ winter season cases (2000-2005), average surface wind speeds at BWI are 55% less than those observed for the remaining winter season cases. The greater frequency of stagnation in winds aloft during winter season high $PM_{2.5}$ cases was previously remarked upon in the discussion of back trajectories (*Figure 26* and *Figure 28*).

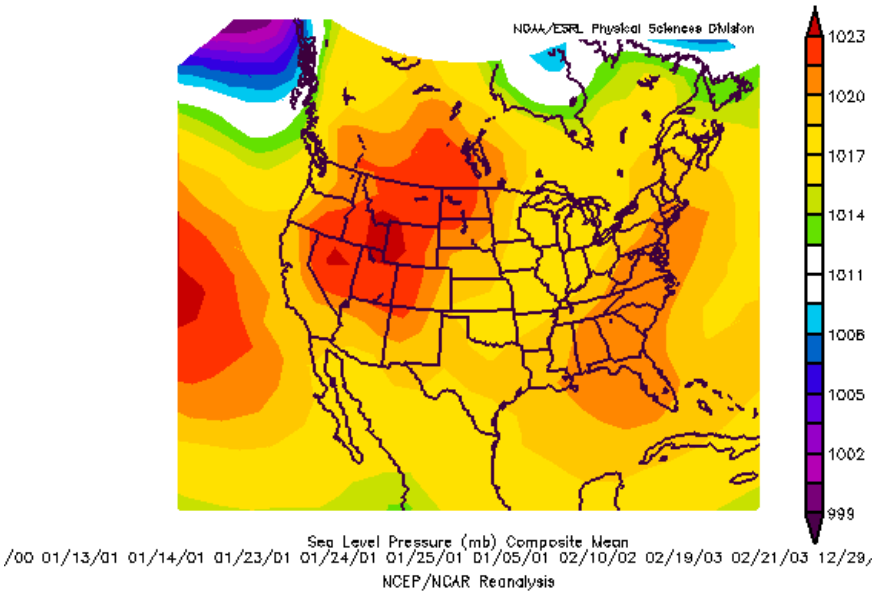


Figure 38. Composite sea level pressure map for all winter season days in Maryland with observed pm $\geq$ 40 μgm^{-3} for the period 2000-2005. Figure courtesy of NOAA/ERSL (http://www.cdc.noaa.gov/cgi-bin/PublicData/getpage.pl).

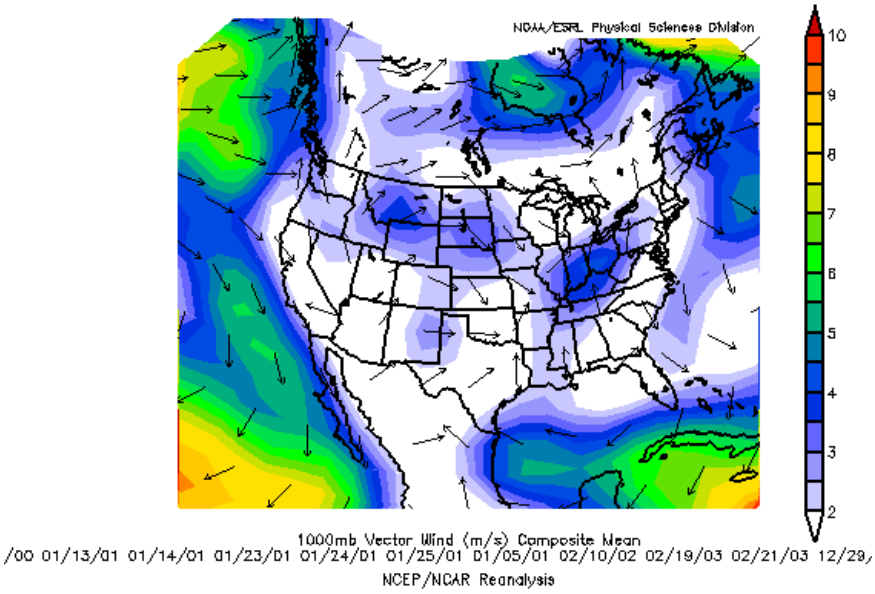


Figure 39. As in Figure 39 but for 1000 mb mean wind speed.

Another important feature of winter season high $PM_{2.5}$ cases is the presence of a very strong near surface inversion. In the summer months, full sun and attendant strong heating of the earth's surface allows deep vertical mixing beginning during the late morning hours. The air mass mixed downward, as discussed above, is not always clean and, in fact, may be laden with sulfate in the "right" (westerly transport, high humidity) conditions. *Figure 23* showed that, on average, in high summer $PM_{2.5}$ cases, mixing tends *not* to lower $PM_{2.5}$ concentrations. In winter, however, while the air mass aloft is typically cleaner than near-surface air parcels, vertical mixing is often quite limited. As a result, although cleaner is mixed downward during the day, it occurs later and often for a very short period. Vertical mixing in winter is limited by weaker solar isolation and can be further reduced in the presence of snow cover. Snow cover, highly reflective, keeps the surface cool and, if warm air advection is occurring above the surface, leads to strong and long lasting morning inversions. This winter season effect is also shown in *Figure 23*. In winter, $PM_{2.5}$ tends to peak in the late morning and early afternoon hours, indicative of limited mixing. Concentrations fall off as mixing finally occurs but only for a few hours until the sun sets and mixing ceases.

Overall, the combination of light winds and limited vertical mixing during winter season events can lead to a more local scale "signature" of pollutants. During summer episodes, sulfate is typically the dominant pollutant, but in winter, organic carbon and nitrates, characteristic of motor vehicle emissions and home heating, become more important. The mix of local and regional effects can vary from episode to episode. These differences are highlighted in the discussion of two winter season episodes that follow.

February 19-21, 2003

This episode followed a coastal storm and heavy snow event on February 16-17. Highest $PM_{2.5}$ concentrations occurred during the period February 19-21 (*Figure 40*). This episode is characterized by a stable air mass, extended periods of near surface stagnation, and transport from the west to southwest. The stability of the boundary layer is, in part, due to a significant snow pack that followed the coastal storm and heavy snow event.

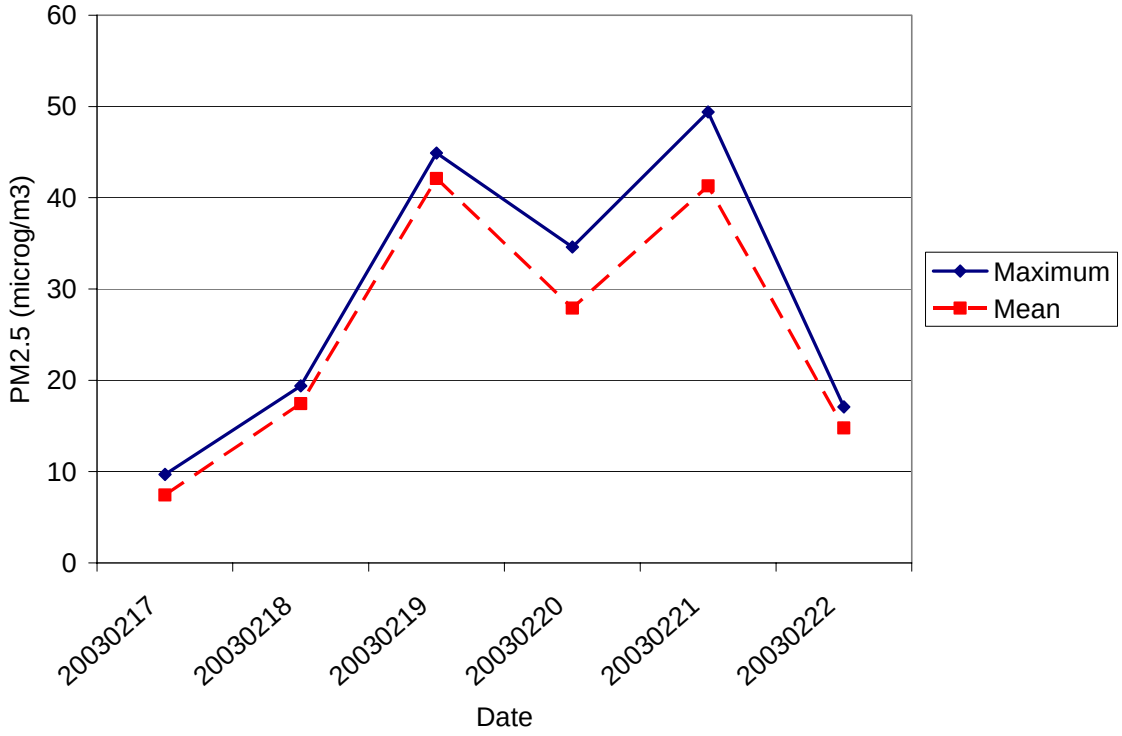


Figure 40. Daily maximum (solid blue line) and mean (dashed red line) $PM_{2.5}$ concentrations for the State of Maryland for the period February 17-22, 2003.

As is common in most winter season $PM_{2.5}$ cases (*Figure 38*), surface high pressure is overhead (*Figure 41*). The snow cover, noted above, results in a very strong surface-based inversion (*Figure 42*). A strong inversion and light winds leads to a typical winter season $PM_{2.5}$ diurnal profile (*Figure 43*) with enhanced morning rush hour concentrations lasting until noon.

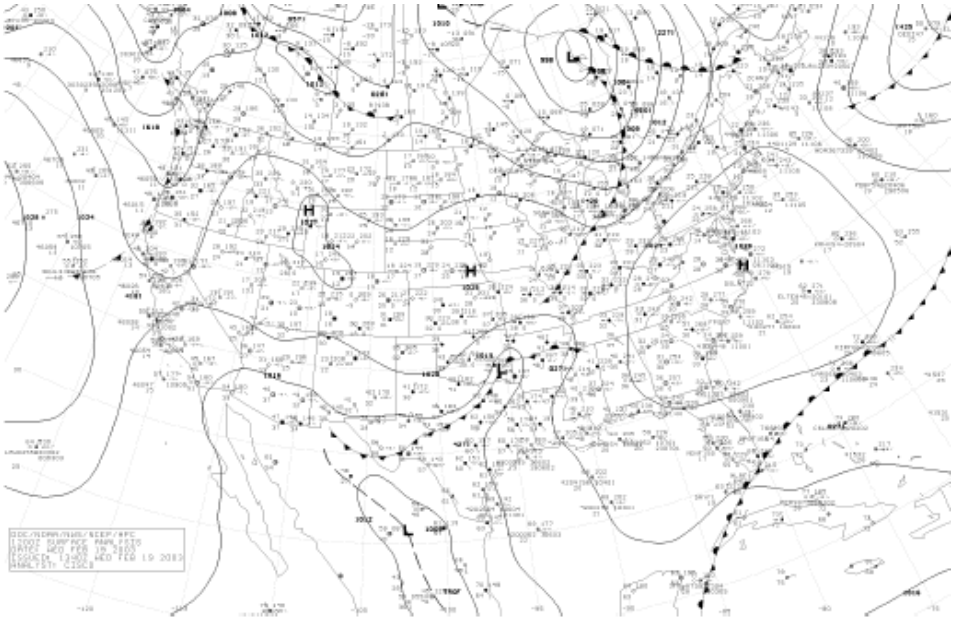


Figure 41. NCEP surface analysis for 1200 UTC on February 19. 2003.

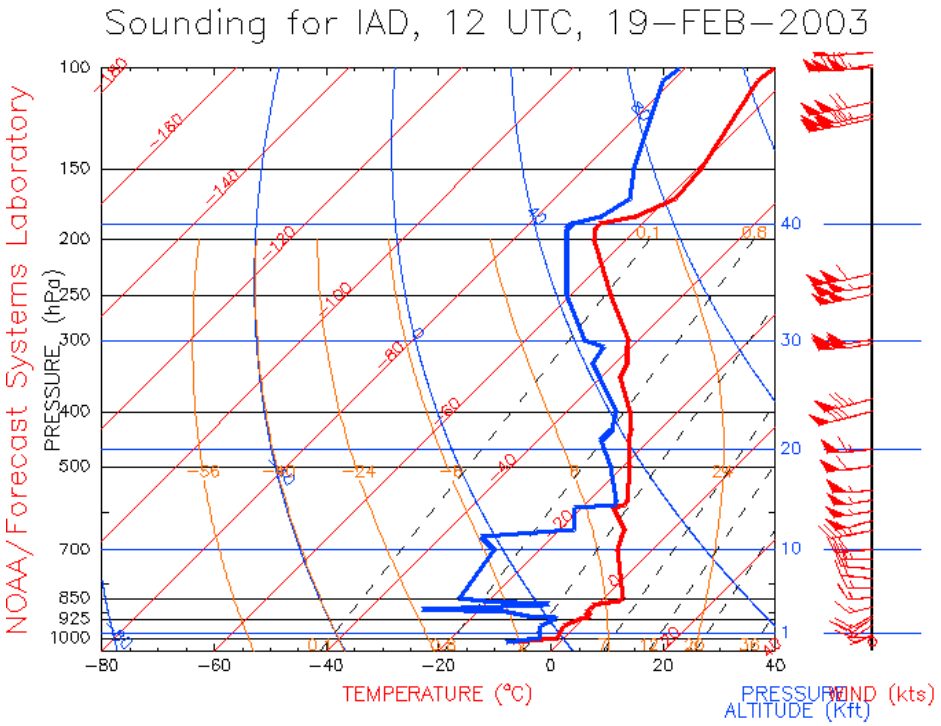


Figure 42. Skew-T diagram giving vertical temperature (red line) and dew point temperature (blue line) observations from the 1200 UTC balloon ascent at Dulles International Airport (IAD) on February 19, 2003. Winds are given in the right margin. Each half-barb is 5 knots, solid pennants are 50 knots.

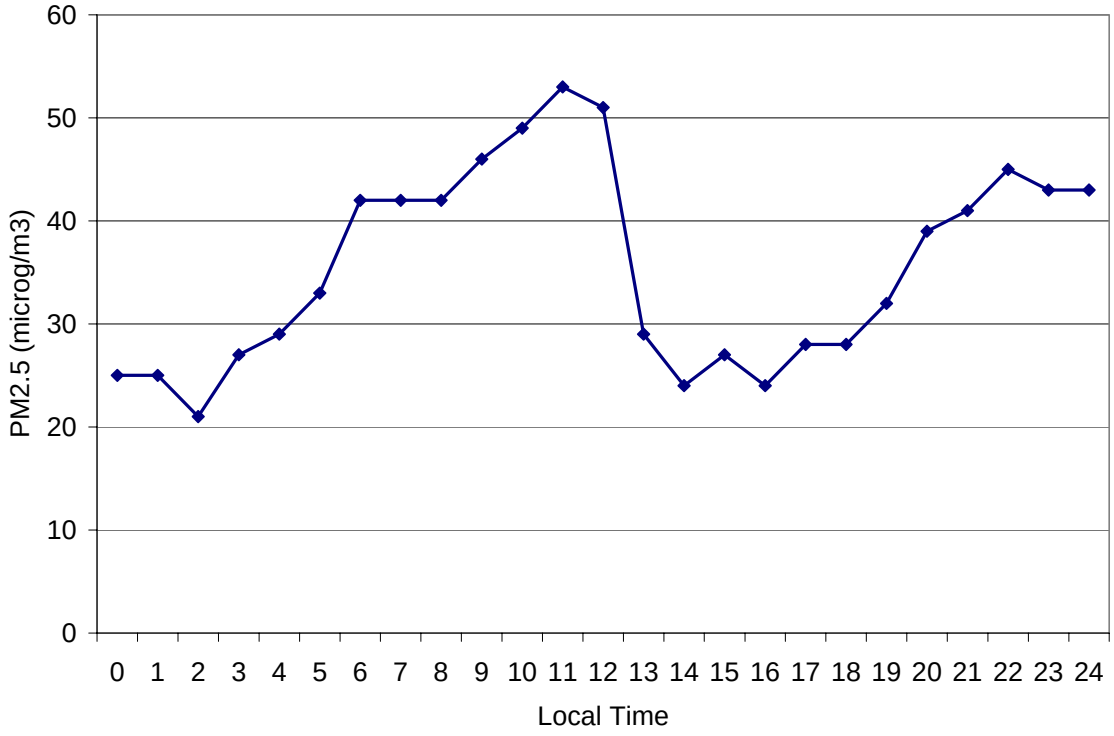


Figure 43. Hourly $PM_{2.5}$ concentrations at Old Town, Maryland for February 19, 2003.

High pressure overhead results in very light winds at BWI with average daily wind speeds only 34% of the winter average. Light surface winds tends to allow local pollutants to accumulate. In addition, relative humidity averaged 85% during the heart of the episode (45% higher than average winter conditions). High relative humidity and cool weather makes for very efficient conversion of NO_x to particulate nitrate. With significant local motor vehicle and home heating sources near Baltimore, additional increases in $PM_{2.5}$ from nitrate is expected. In addition, westerly to southwesterly transport was a factor throughout the episode (*Figure 44*). Westerly transport is characterized by a strong sulfate component. As a result, this episode is a combination of both local and regional impacts. Speciated $PM_{2.5}$ measurements are available at Essex and Washington DC only for February 20th. All $PM_{2.5}$ constituent are enhanced during this event in the range of 2-7 μgm^{-3} above their winter season averages. Relative to winter average concentrations, SO_4 shows the biggest relative increase – three times normal – although absolute concentrations are only 9.7 μgm^{-3}. Organic carbon (OC) is the largest absolute contributor to $PM_{2.5}$ on this day with concentrations (depending on the k factor applied) of 13-17 μgm^{-3}. NO_3 is also enhanced and reaches 4.8 μgm^{-3}. Observations in Philadelphia show similar concentrations but with a stronger NO_3 component (8.6 μgm^{-3}) with OC again forming the largest part of the observed $PM_{2.5}$. Concentrations from the IMPROVE monitor in Washington DC for the entire year are

shown in *Figure 45*. During this period, and through the winter months, the highest $PM_{2.5}$ cases, marked "W" for the worst air quality days, is a combination of enhanced sulfate and nitrate.

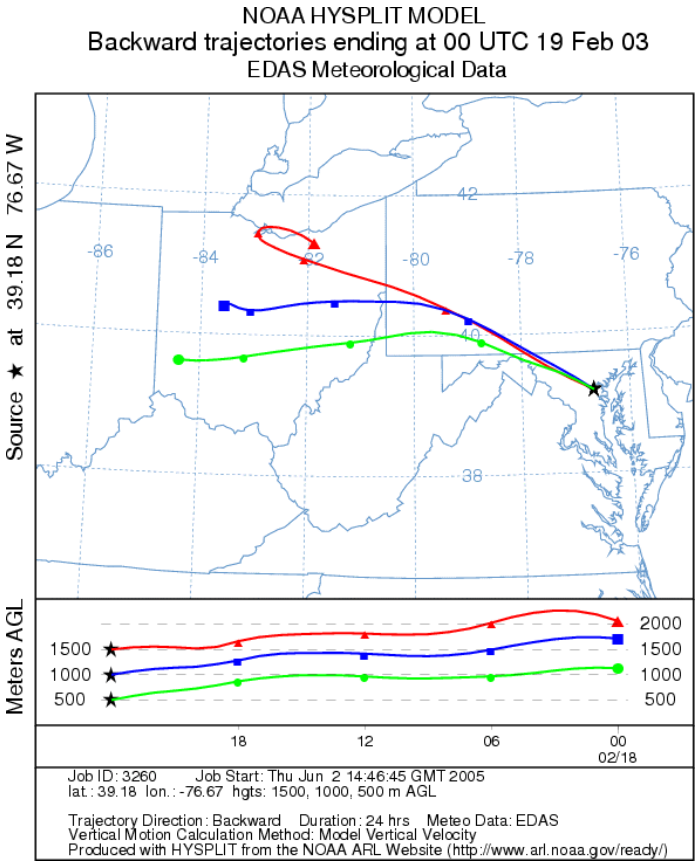


Figure 44. HYSPLIT back trajectories terminating at BWI for 0000 UTC on February 19, 2003. Trajectories are for 24-hours of travel and terminate at 500 (green line), 1000 (blue line) and 1500 (red line) meters agl.

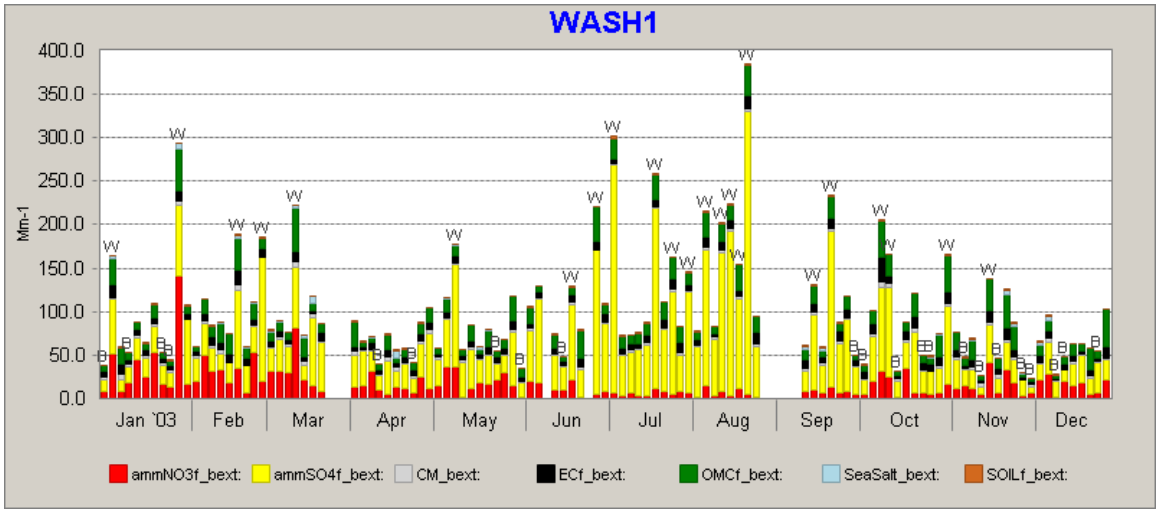


Figure 45. Daily speciated $PM_{2.5}$ concentrations from the Washington, DC IMPROVE monitor for 2003. Labels "W" and "B" represent the worst and best air quality days during the year.

January 12-15, 2001

The highest $PM_{2.5}$ concentrations during this episode were observed in Maryland from January 12-15 while, region-wide, the peak of the episode occurs on January 14 (*Figure 46*). In the synoptic scale, the episode is characterized by surface high pressure over the region consistent with the winter average conditions (cf. *Figure 38* and *Figure 47*). The keys to the episode are westerly transport in the onset stage, followed by re-circulation and moistening of the air mass in the mature stage.

The onset of the episode was characterized again by westerly transport (*Figure 48*). By the morning of the 13th, $PM_{2.5}$ is strongly enhanced regionwide (*Figure 49*) and a strong morning inversion keeps concentrations high until noon. By January 14th, stagnation becomes pronounced. The 48-hour back trajectory in *Figure 50* suggests that the air mass advected from the west on January 12th lingers over Maryland through the 14th. Surface weather observations at BWI show the extent of the stagnation. While average winds during the afternoon hours on January 12th are ~ 6 kts, they fall off to 2 kts or less from the 13th through the 15th – the period of high $PM_{2.5}$ concentrations. In addition, relative humidity, which favors nitrate and sulfate formation processes, increases through the period as well (*Figure 51*). A strong surface based inversion is again present (*Figure 52*). While no STN data is available for this period, IMPROVE monitor in Washington DC does provide data for January 13 (*Figure 53*). $PM_{2.5}$ on January 13 is the highest concentrations during the winter monthis in Figure 54 and the large nitrate contribution is consistent with high humidity, stagnation, local scale event driven by motor vehicle and home heating emissions.

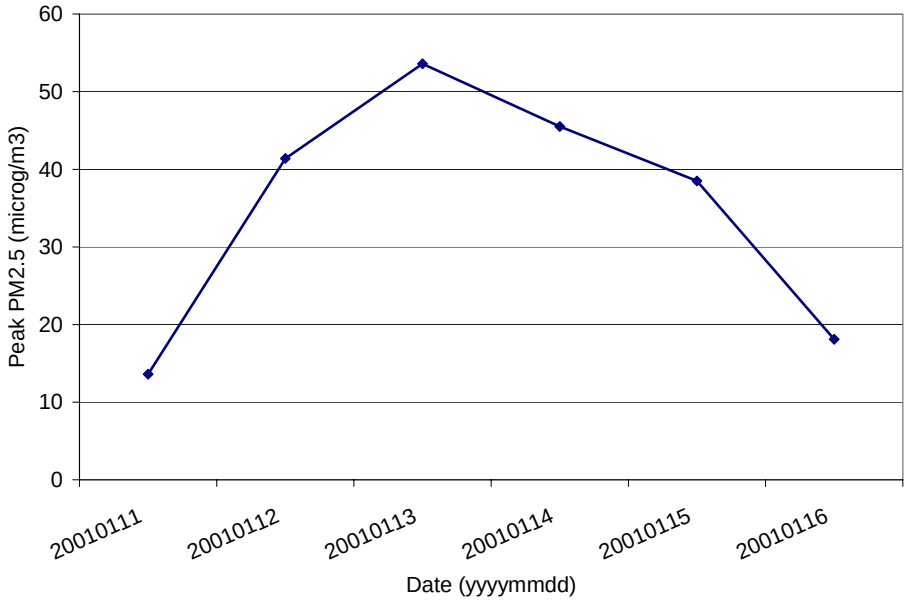


Figure 46. Daily maximum $PM_{2.5}$ concentrations for the State of Maryland for the period January 11-16, 2001.

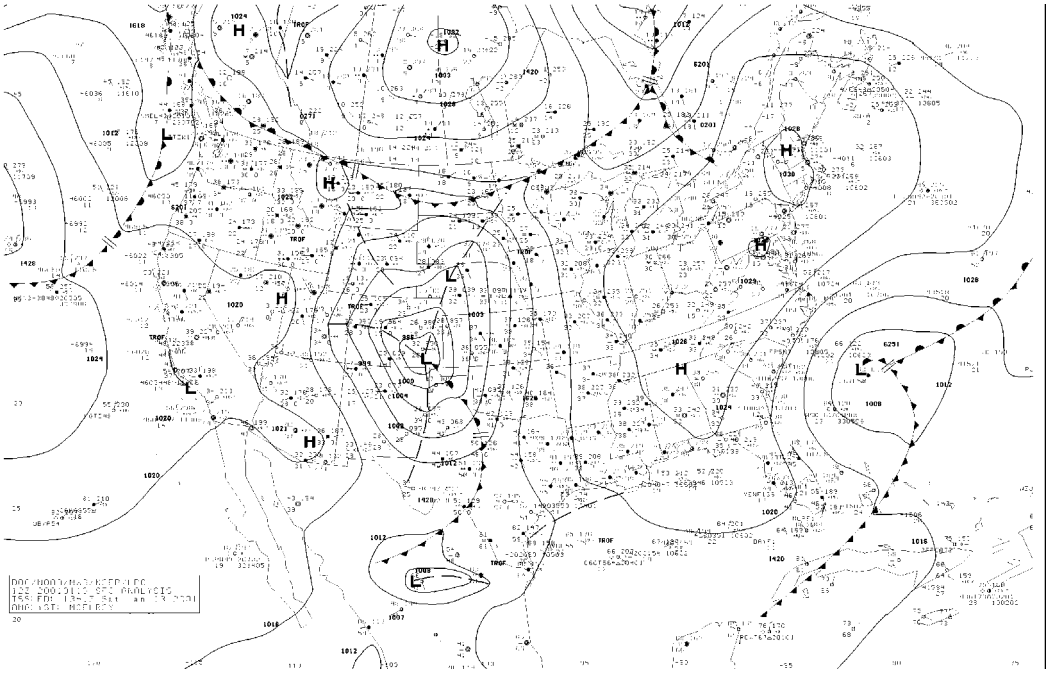


*Figure 47***.** NCEP surface analysis for 1200 UTC on January 13. 2001.

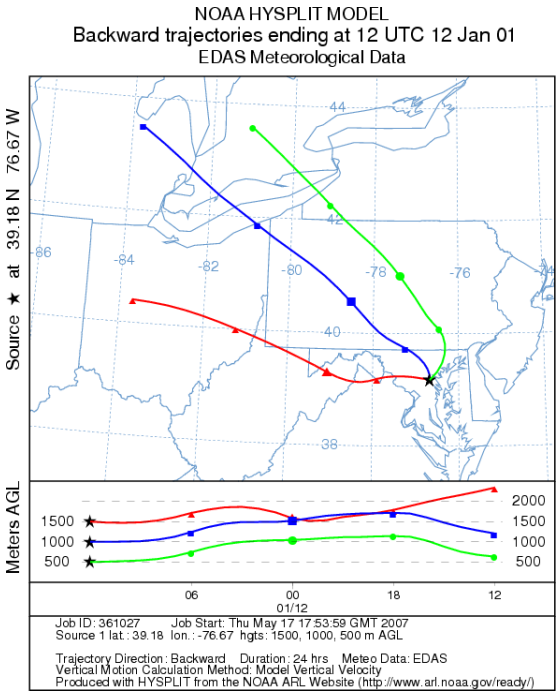


Figure 48. HYSPLIT back trajectories terminating at BWI for 1200 UTC on January 12, 2001. Trajectories are for 24 hours of travel and terminate at 500 (green line), 1000 (blue line) and 1500 (red line) meters agl.

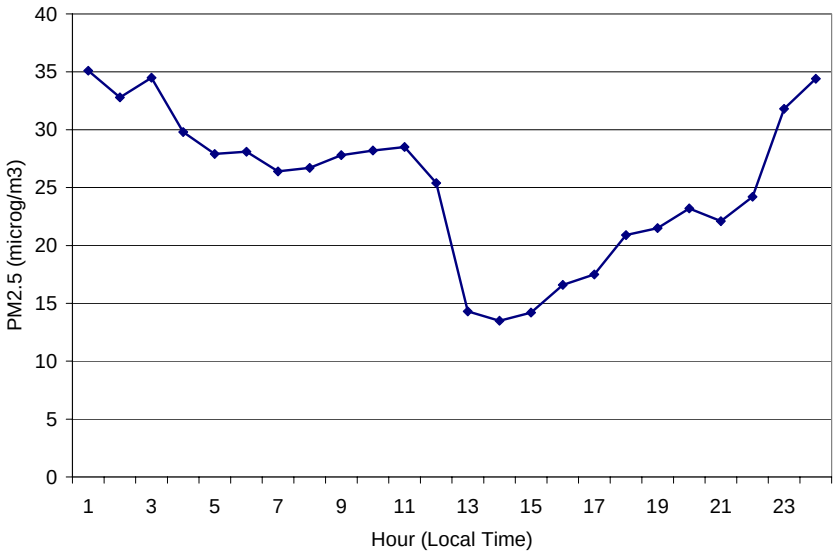


Figure 49. Hourly $PM_{2.5}$ concentrations at Old Town, Maryland for January 13, 2001.

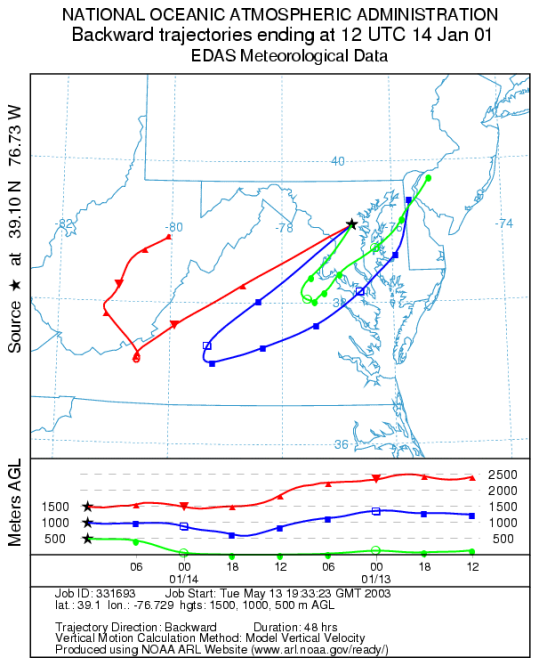


Figure 50. HYSPLIT back trajectories terminating at BWO for 0000 UTC on January 14, 2001. Trajectories are for 48 hours of travel and terminate at 500 (green line), 1000 (blue line) and 1500 (red line) meters agl.

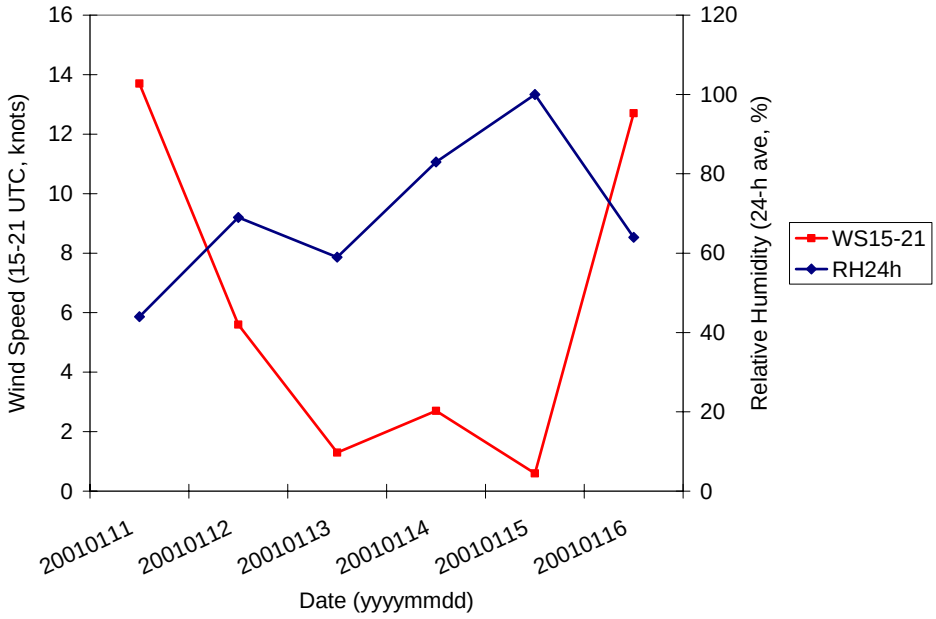


Figure 51. Daily average relative humidity (%) at BWI (red line, right axis) and mid-day (1500-2100 UTC) average wind speed for January 11-16, 2001.

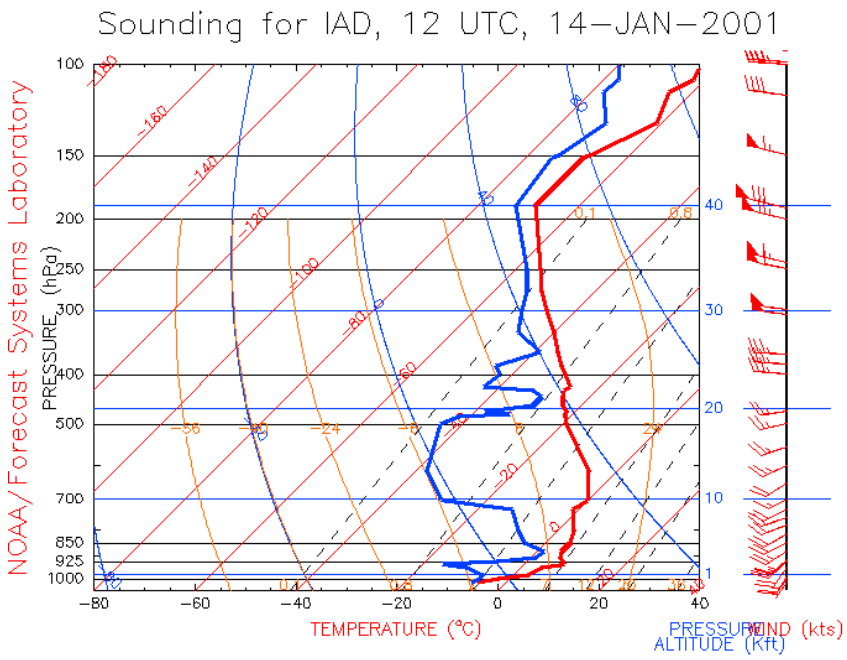


Figure 52. Skew-T diagram giving vertical temperature (red line) and dew point temperature (blue line) observations from the 1200 UTC balloon ascent at Dulles International Airport (IAD) on January 14, 2001. Winds are given in the right margin. Each half-barb is 5 knots, solid pennants are 50 knots.

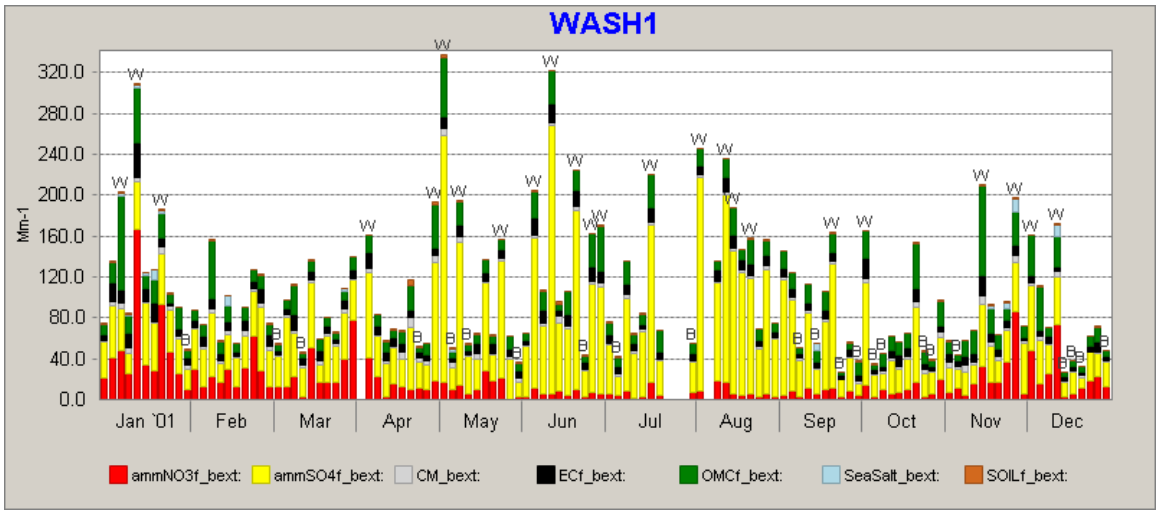


Figure 53. Daily speciated $PM_{2.5}$ concentrations from the Washington, DC IMPROVE monitor for 2001. Labels "W" and "B" represent the worst and best air quality days during the year.

Conclusions

This report summarizes observations of fine particulate matter with an aerodynamic radius of ≤ 2.5 μm, also known as $PM_{2.5}$, in the State of Maryland during the period 2000-2005. Average $PM_{2.5}$ concentrations for this period at monitors across Maryland range from 12-17 μgm^{-3} with the highest concentrations observed at urban scale sites. Although urban monitors observe the highest concentrations, all Maryland $PM_{2.5}$ monitors are strongly correlated and the correlation between monitors is highest during the summer season.

The data summary indicates that $PM_{2.5}$ is ubiquitous in Maryland, is not an overwhelmingly urban pollutant, and that all locations share in a common, regional scale "load" of $PM_{2.5}$. Although there is no fool-proof method to quantify the regional scale "load" of $PM_{2.5}$, comparisons of urban, suburban and remote rural monitors suggest that the regional component of $PM_{2.5}$ accounts for roughly 60-75% of the total observed $PM_{2.5}$. This fraction increases to 80-90% during the summer season.

$PM_{2.5}$ concentrations peak during the summer season (June-August) in Maryland although urban scale sites also have a secondary maximum during the winter (December-February) months. The summer maximum are driven primarily by increases in the amount of sulfate while winter season peaks are driven more by increases in nitrogen and carbon compounds. $PM_{2.5}$ concentrations also vary by the day of the week, on the order of 2-3 μgm^{-3}, with highest concentrations occurring near the end of the work week and lowest concentrations on Sunday. This reflects day of week differences in motor vehicle and industrial emissions. While average concentrations do not vary significantly by the day of the week, the frequency of high $PM_{2.5}$ concentrations days (90^{th} percentile) is much greater during the work week.

$PM_{2.5}$ concentrations have a daily (diurnal) cycle with highest concentrations during the morning and afternoon rush hours, when emissions are highest and vertical mixing is weakest, and lowest concentrations during the well-mixed (diluted) afternoon hours. The diurnal cycle is markedly different for the most severe (90^{th} percentile) cases. In those cases, the mid-day dilution effects are less evident so that concentrations remain nearly unchanged through the daylight hours. This effect is more pronounced in the summer months and suggests that the air aloft, which mixes downward in the afternoon, is heavily laden with transported $PM_{2.5}$. The highest $PM_{2.5}$ cases are characterized overwhelming by westerly transport of air parcels although, in winter, there is a secondary maximum of cases where re-circulation, or stagnation, occurs. Observations at rural monitors west of Maryland show that, on the worst $PM_{2.5}$ days, this air mass is primarily made up of sulfate particles.

Appendix A: Assessment of FRM Data

Compliance with the $PM_{2.5}$ NAAQS is based on observations made by gravimetric filters using the Federal Reference Method (FRM). Due to cost constraints, only a fraction (~ 20%) of all FRM monitors report daily with the remainder reporting every third day. For Maryland, typically 10-17 FRM monitors report every third day with only 3 monitors (all near Baltimore) reporting daily. This discontinuity in sampling frequency may introduce problems in statistical analyses – particularly in the analysis of statewide concentrations and the analysis of episodic high $PM_{2.5}$ episodes. Overall, the high sampling frequency cases tend to observe higher $PM_{2.5}$ concentrations (Table A1) suggesting that the data collected on these days are qualitatively different from the low sampling days. This task will assess whether there are statistically significant differences between the high sampling frequency (3rd day) and low sampling frequency (days 1 and 2) cases.

Complicating any analysis of $PM_{2.5}$ data, particularly daily peak concentrations, is the non-normal distribution of the observed data. As shown in Figure A1, the distribution of daily peak 24-hour average $PM_{2.5}$ concentrations in Maryland, as well as 24-hour concentrations at any single monitor within the state, is highly skewed. Basic statistical measures for statewide maximum $PM_{2.5}$ and a Baltimore monitor (Old Town) are given in Table A1. The large difference between mean (17.7 μgm^{-3}) and median (15.5 μgm^{-3}) concentrations for the statewide maximum concentrations is an initial clue that the distribution is not normal. Skewness and kurtosis measures are generally used to diagnose differences from normal in any given distribution. *Skewness* is a measure of the symmetry of the distribution about its mean. If skewness is significantly non-zero, the distribution is asymmetric. A standard threshold above which the skewness coefficient is considered significant is if the absolute value of the skewness coefficient divided by the standard error of skewness (SES) is > 2. The SES is typically defined as $(6/n)^{1/2}$, where *n* is the number of cases in the sample. For both statewide monitors and Old Town, and all other monitors as well, the absolute values of skewness/SES are well in excess of 2 and thus the distribution is asymmetric. The positive value for skewness means that the distribution is dominated by a strong "right tail" of high $PM_{2.5}$ concentrations.[1] *Kurtosis* is a measure of whether the center of the distribution is strongly "peaked". Positive kurtosis means that the distribution is more peaked, that is, less smooth and bell shaped, than the normal distribution. The kurtosis coefficient is considered significant, and the sample distribution significantly different from normal, if the measure of kurtosis divided by the standard error of kurtosis (SEK) is > 2. The SES is defined as $(24/n)^{1/2}$. In this case, the kurtosis measure is significantly positive meaning the sample distribution is more centrally peaked than normal. In summary, the statewide maximum $PM_{2.5}$ distribution, and individual monitors as well, are significantly different from normal. The distribution is asymmetric, with a strong right tail, and strongly peaked.

As an aside, Figure A2 shows that a natural log-transform of the $PM_{2.5}$ data is sufficient to approximate a normal distribution. To the extent that this dataset is used for $PM_{2.5}$ forecast algorithm development using linear regression techniques, a log transform of the predicand is necessary for best results.

The results above raise a number of issues. First, they call into question the usefulness of mean concentrations as the NAAQS for $PM_{2.5}$. The NAAQS is based on yearly mean measures yet, in a highly skewed distribution, mean concentrations tell little about the nature of the observed distribution. In this case, the mean value is strongly affected by the high $PM_{2.5}$ outliers and does not accurately reflect the "usual" $PM_{2.5}$ exposure of the population. For example, if the highest 5^{th} percentile (> 37.2 μgm^{-3}) is excluded from the Old Town monitor, mean concentrations fall from 16.7 μgm^{-3} (above the NAAQS of 15.5 μgm^{-3}) to only 14.5 μgm^{-3}. Second, and of more importance for the present task, the usual statistical measures of whether two samples are drawn from the same population are not applicable when the distributions are not normal.

In this case, therefore, the appropriate statistical test must be non-parametric. That is, the test must be applicable without assumptions as to the shape of the distribution. The most common non-parametric test to determine if two samples are from the same distribution is the Komolgorov-Smirnov (K-S) test. The K-S test statistic (D_s) looks for the largest difference, in absolute value, between the distributions of two samples of data. In specific, D_s is the largest difference between the empirical cumulative distribution functions of the two samples. If D_s is sufficiently large, the null hypothesis, that the two samples come from the same distribution, can be rejected. The critical level of D_s depends on sample size (Wilks, 2006). In Figure A3, a rough illustration of the D_s statistic is shown. The largest difference in our data appears to occur in the extreme of the distribution – the highest $PM_{2.5}$ cases.

The D_s statistic is given below for two samples of *n* and *m* observations. The null hypothesis that two samples come from the same distribution is rejected at $\alpha * 100\%$ if:

$$D_s > \left[-\frac{1}{2}\left(\frac{1}{n}+\frac{1}{m}\right)\ln\left(\frac{\alpha}{2}\right)\right]^{1/2}$$

In this case, the small frequency cases (1 and 2 day observations) have 1381 members (*n*) and the large frequency cases (3^{rd} day observations) have 727 members (*m*). Using results from the SYSTAT statistical software package two-sample K-S test, D_s = 0.1018 for these samples and the null hypothesis is accepted at p = 0.0001. As a result, the null hypothesis is accepted at a 99.9% likelihood and the high and low frequency samples can be considered to derive from the same distribution.

Given this result, we are able to make use of the entire statewide dataset of low and high frequency sampling cases in the succeeding analyses. It also gives forecasters a much larger (3x) database from which to train statistical forecast algorithms. In the analyses of statewide $PM_{2.5}$ that follows in Tasks 3 and 4 of this report, the entire daily database will be used except where otherwise noted.

Maryland Statewide Daily Maximum $PM_{2.5}$ 2000-2005		
	High Sampling Frequency (Every 3rd Day)	Low Sampling Frequency (Days 1 and 2)
Median	17.0	14.8
Mean	19.0	17.0
90th %ile	33.3	29.5
95th %ile	39.2	36

Table A1. Basic statistical measures for statewide maximum $PM_{2.5}$ for every third day when all FRM monitors report and other days when only daily FRM monitors report.

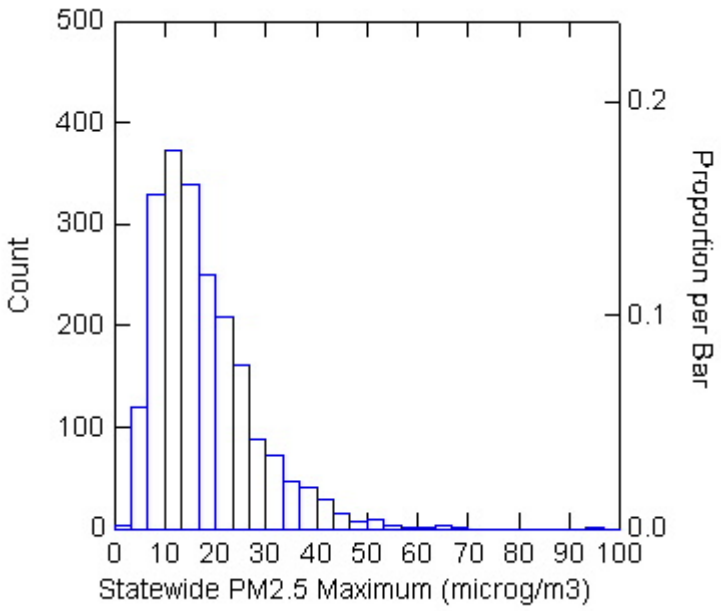


Figure A1. Daily maximum $PM_{2.5}$ concentrations (24-hour average) from all FRM monitors in the State of Maryland.

\

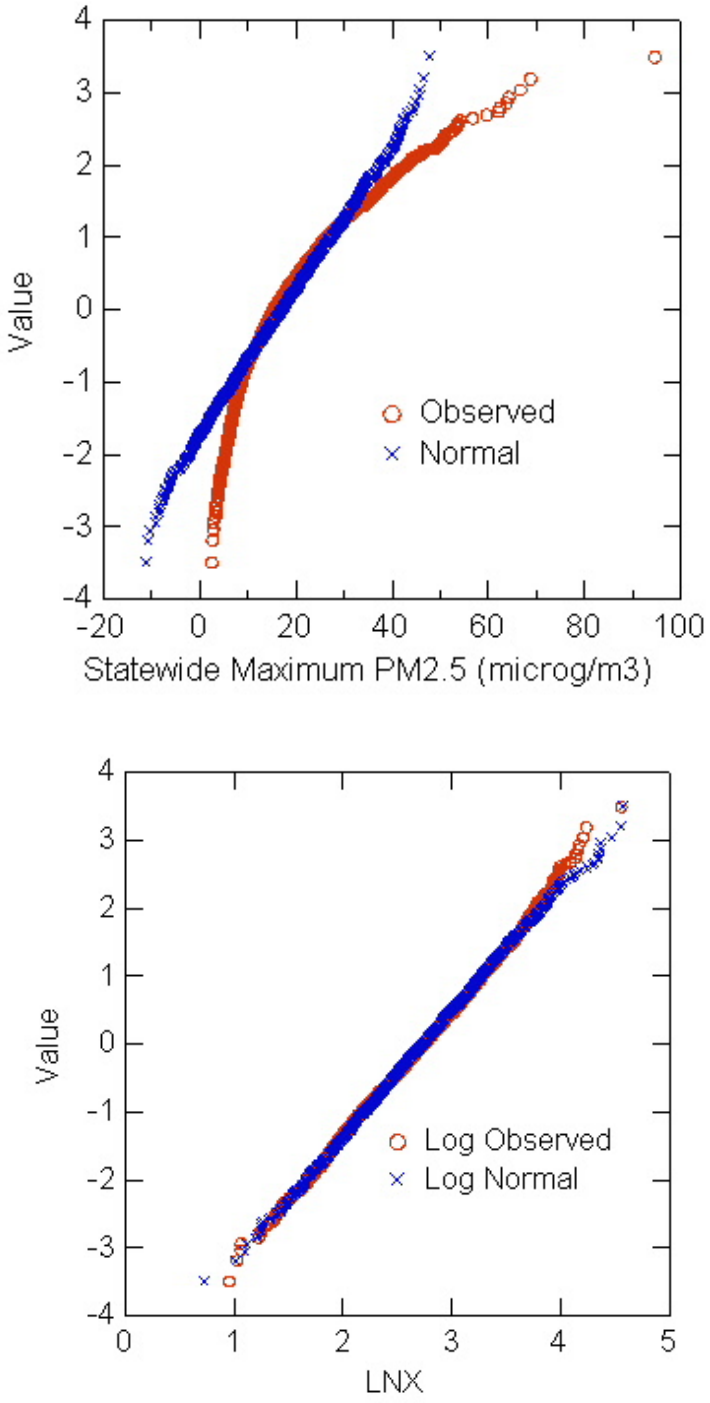


Figure A2. Probability plots for statewide maximum daily $PM_{2.5}$ concentrations and normal distributions given same mean and standard deviation. Top panel: Unmodified statewide $PM_{2.5}$ concentrations. Bottom panel: Statewide $PM_{2.5}$ concentrations

transformed by the natural log function. The log-transformed $PM_{2.5}$ data best approximates a normal distribution.

Maryland $PM_{2.5}$ 2000-2005 Basic Statistical Measures		
Statistical Measure	**Statewide**	**Old Town**
N	2107	1918
Median	15.5	14.7
Mean	17.7	16.7
Standard Deviation	9.8	9.2
Skewness	1.49	1.19
Standard Error of Skewness (SES)	0.05	0.06
Kurtosis	3.78	1.67
Standard Error of Kurtosis (SEK)	0.11	0.11
Skewness/SES	29.8	19.8
Kurtosis/SEK	34.4	15.2

Table A2. Statistical summary for statewide daily maximum $PM_{2.5}$ and for the monitor at Old Town in Baltimore.

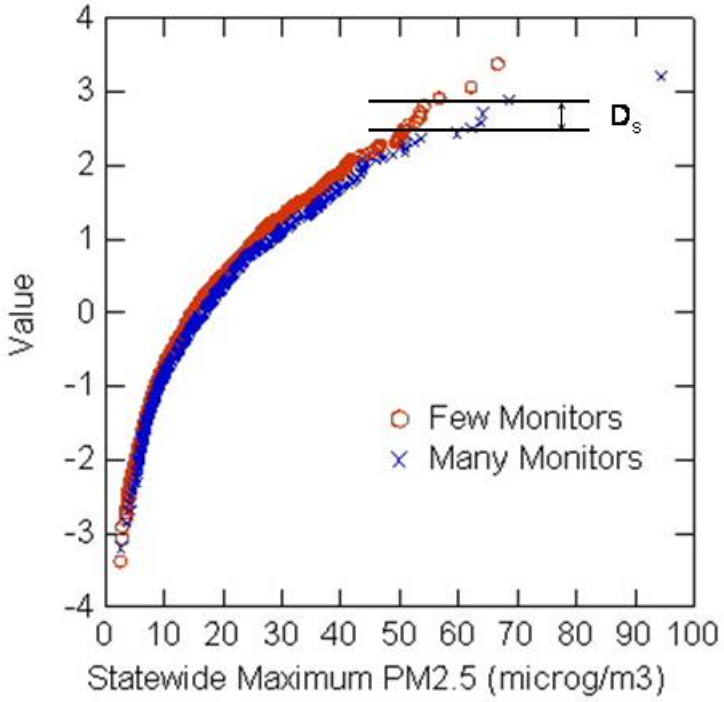


Figure A3. Example of the D_s measure used in the K-S two sample test.

Footnotes

1. The conventional moments based measure of symmetry, the skewness coefficient, is based on the cube of the difference between any data point and the mean. As a result, the measure is criticized as neither statistically robust nor resistant. The measure is, in particular, sensitive to the influence of outliers due to the cubic function. An alternative measure of symmetry that is more resistant is the Yule-Kendall Index:

$$\gamma_{YK} = (q_{0.25} - 2* q_{0.25} + q_{0.25})/IQR$$

Where q_n is the nth percentile of the data and IQR (inter-quartile range) is the difference between the 75^{th} and 25^{th} percentile.

In this case, $\gamma_{YK} = 11.9$ which corroborates the skewness coefficient results above that concluded that the sample distribution was asymmetric.

References

Wilks, D.S., 2006: *Statistical Methods in the Atmospheric Sciences* (2^{nd} Ed.), Elsevier Press.

SYSTAT, 2002: *SYSTAT 10.2, Statistics II*, SYSTAT Software, Inc, www.systat.com.

Appendix B: FRM Stations

Rural Monitors AQS Code	Locations	State
240430009	Hagerstown	MD
420990301	Little Buffalo SP	PA
420270100	PSU Arboretum	PA
511390004	Luray	VA
540890001	Keeney Knob	WV

Maryland Monitors

240030014	Davidsonville
240251001	Edgewood
240053001	Essex
240150003	Fair Hill
245100035	FMC-Curtis Bay
245100006	North East Police
245100007	North West Police
245100008	South East Police (6/01 - current)
245100049	Westport Elem
240031003	Glen Burnie
240430009	Hagerstown
245100040	Old Town
240051007	Padonia
240032002	Riviera Beach
240313001	Rockville
240330030	HU-Beltsville (8/04 > current)
240338003	Prince Geo. Equest. (5/02 - current)
240330002	Greenbelt GSFC (7/02-4/04)

BCC = Baltimore (Urban) Monitors

240053001	Essex
245100006	North East Police
245100007	North West Police
245100008	South East Police (6/01 - current)
245100035	FMC-Curtis Bay
245100049	Westport Elem
245100040	Old Town

BFA = Baltimore Forecast Area

240053001	Essex
245100006	North East Police
245100007	North West Police
245100008	South East Police (6/01 - current)

245100035	FMC-Curtis Bay
245100049	Westport Elem
245100040	Old Town
240030014	Davidsonville
240251001	Edgewood
240150003	Fair Hill
240031003	Glen Burnie
240051007	Padonia
240032002	Riviera Beach

BMA = Baltimore Metropolitan Area

240053001	Essex
245100006	North East Police
245100007	North West Police
245100008	South East Police (6/01 - current)
245100035	FMC-Curtis Bay
245100049	Westport Elem
245100040	Old Town
240251001	Edgewood
240150003	Fair Hill
240031003	Glen Burnie
240051007	Padonia
240032002	Riviera Beach

DCSUB = Washington DC Suburban

240030014	Davidsonville
240313001	Rockville
240330030	HU-Beltsville (8/04 > current)
240338003	Prince Geo. Equest. (5/02 - current)
240330002	Greenbelt GSFC (7/02-4/04)

MDSUB = Maryland Suburban

240030014	Davidsonville
240313001	Rockville
240330030	HU-Beltsville (8/04 > current)
240338003	Prince Geo. Equest. (5/02 - current)
240330002	Greenbelt GSFC (7/02-4/04)
240251001	Edgewood
240150003	Fair Hill
240051007	Padonia
240032002	Riviera Beach
240031003	Glen Burnie

Appendix C: Location of Selected Maryland FRM Stations

This document is in Power Point format and can be accessed at:
http://www.meteo.psu.edu/~wfryan/mde/frm-monitors.ppt

Appendix D: Episode Selection

Episodes Selected

January 12-15, 2001
June 26-30, 2001
August 5-10, 2001
July 18-22, 2002
February 19-21, 2003
June 25-30, 2003
October 8-10, 2003
July 20-22, 2003
August 11-14, 2005

Criteria:

1. No more than 10 episodes in initial selection.
2. Episode spans ≥ 3 days.
3. Episode days must be within the 90th percentile of the maximum and mean $PM_{2.5}$ distributions statewide.
4. Episodes do not include major (fireworks) holidays of New Years and July 4th.
5. Single monitor Code Orange peak concentrations must occur on at least 2 days during episode.
6. More than 2 monitors must exceed the Code Orange threshold during the episode.

Discussion of Criteria

Criteria 1

A full analysis of multi-day pollution episodes are time consuming. Criteria 1 sets a limit on the number of episodes that is proportional to the effort that can be applied. In this case, nine are selected. As is often the case, full data is sometimes not available to analyze each possible episode so that the final number of episodes analyzed will likely be somewhat smaller - in the range of 7-8.

Criteria 2

As will be noted in more detail in the final report, the bulk of all high $PM_{2.5}$ episodes occur in multi-day episodes. This reflects, in part, the dependence of $PM_{2.5}$ on meteorological conditions. In addition, single day "spikes" in $PM_{2.5}$ concentrations can often be due to local scales effects that are difficult to analyze at the resolution available from archived materials. This also helps to remove cases that are dominated by mainly local scale effects. Longer time scale events can be influenced by synoptic scale (2-5 day)

weather patterns that are well resolved by archived weather data. Finally, making the episodes extend at least three days makes sure that at least one large sampling frequency day (every third day is a large sample observation) is included. This is particularly true in Maryland as all the daily monitors are grouped close to center city Baltimore and may not always reflect regional conditions.

Criteria 3

The days in the selected episodes must be within the 90^{th} percentile of the maximum and mean concentrations statewide. Because of the differences in sampling frequency, this criteria is relaxed for maximum concentrations so that if one day within an episode does not reach the 90^{th} percentile but is bounded on both sides by cases which do reach the criteria, the episode is still considered. For example, if July 17, 19, 20 and 21 are above the 90^{th} percentile, and the 18^{th} is not, the episode is considered to extend from July 17-21 as a 4 day episode. For the episodes selected above, 37 of 39 days were within the 90^{th} percentile of the mean statewide concentrations.

Criteria 4

Some of the highest $PM_{2.5}$ cases occur on days of fireworks displays (December 31-January 1, July 4). Although multi-day episodes may still occur that include one of these days and, therefore, not totally influenced by fireworks, it is difficult to separate the different effects and so these cases have been excluded.

Criteria 5-6

A fairly large (~ 16) set of episodes met Criteria 1-4, so that Criteria 5-6 were added to limit the number of episodes. Both Criteria 5 and 6 add a short term severity measure to the selection that is otherwise dominated by longer term, regional effects.