Regional Aerosol Intensive Network (RAIN)

Preliminary Data Analysis

Prepared by NESCAUM

For the Mid-Atlantic/Northeast Visibility Union (MANE-VU) Regional Planning Organization



May, 2006

Members of Northeast States for Coordinated Air Use Management

Arthur Marin, Executive Director Northeast States for Coordinated Air Use Management

Anne Gobin, Bureau Chief Connecticut Department of Environmental Protection, Bureau of Air Management

James P. Brooks, Bureau Director Maine Department of Environmental Protection, Bureau of Air Quality

Barbara Kwetz, Director Massachusetts Department of Environmental Protection, Bureau of Waste Prevention

Robert Scott, Director New Hampshire Department of Environmental Services, Air Resources Division

William O'Sullivan, Director New Jersey Department of Environmental Protection, Office of Air Quality Management

David Shaw, Director New York Department of Environmental Conservation, Division of Air Resources

Stephen Majkut, Chief Rhode Island Department of Environmental Management, Office of Air Resources

Richard A. Valentinetti, Director Vermont Department of Environmental Conservation, Air Pollution Control Division

Regional Aerosol Intensive Network (RAIN)

Preliminary Data Analysis

Prepared by NESCAUM

for the Mid-Atlantic/Northeast Visibility Union (MANE-VU) Regional Planning Organization

May, 2006

Submitted to the United States Environmental Protection Agency, Region III in partial fulfillment of requirements for EPA agreement XA-97318101-0 to the Ozone Transport Commission

Project Director

Gary Kleiman, NESCAUM

Editors

Paul Miller, NESCAUM Gary Kleiman, NESCAUM

Principle Contributors

Alan R. Leston, NESCAUM George Allen, NESCAUM

US EPA Project Officer

Marcia Spink, US EPA Region III

Acknowledgments

Funding support for the RAIN network instruments and coordination was provided by the U.S. Environmental Protection Agency through their support of the MANE-VU Regional Planning Organization (Grant # XA-97318101-0). Maintenance and operation of the sites was provided in-kind by the individual state air quality agencies where the sites are located. NESCAUM would like to thank the following individuals whose contributions and efforts greatly assisted in this work:

Ted Erdman	Technical Program Manager, U.S. EPA RIII
Norm Beloin	Air Monitoring Team Leader U.S. EPA RI
Paul Norton	Head of Field Operations, CT DEP
David Krask	Division Chief, Ambient Air Monitoring, MDE
Chris Smith	Section Head, Technical Support and Special Projects, MDE
Andy Johnson	Laboratory & Quality Assurance Coordinator, ME DEP
Jeff Emery	Ambient Monitoring Coordinator, ME DEP
Reid Matusek	Sr. Air Pollution Control Engineer, CT DEP
Peter Babich	Environmental Analyst III, CT DEP
Adam Augustine	Mohawk Mt. Site Operator, CT DEP
Janine McKnight	Piney Run Site Operator, Univ. of Maryland, Frostburg
John Lane	Site Operator, ME DEP
Paul Nichols	Site Operator, ME DEP
Don Darling	Eastern Maine Field Monitoring Supervisor, ME DEP

Printed: May 2006

Acknowl	edgments	iii	
Executive	e Summary	v	
1. Intro	duction to the Rural Aerosol Intensive Network (RAIN)	1-1	
2. Con	parative Data Analysis		
2.1.	RAIN Sulfate Data Analysis		
2.2.	IMPROVE SO ₄ Comparison		
2.3.	Conclusions		
3. Carb	oon Data Analysis		
3.1.	IMPROVE Seasonality		
3.2.	Temperature Dependence		
3.3.	Wood Smoke Impact		
3.4.	Regional Organic Carbon		
3.5.	Semi-Continuous Organic Carbon vs. Ozone		
3.6.	PAMS Data Comparison		
3.7.	Conclusions		
4. Reco	onstruction of Fine Mass		
4.1.	McFarland Hill, Maine		
4.2.	Mohawk Mountain, Connecticut		
4.3.	Piney Run, Maryland		
4.4.	Conclusions		
5. Visi	bility Reconstruction		
5.1.	Approach and Discussion		
5.2.	Conclusions		
6. Refe	rences		
Appendix	A: RAIN Instrumentation		
Appendix	Appendix B: Data Validation and Caveats		
Appendix	C: Sulfate Monitor Performance Issues	C-2	

Table of Contents

Executive Summary

This technical memo focuses on the air monitoring aspect of regional haze planning. Although the Interagency Monitoring of Protected Visual Environments (IMPROVE) network currently provides the core data for the Regional Haze program we need more information to successfully achieve visibility planning goals. Additional data will be critical for understanding the complex issues associated with the organic carbon portion of fine aerosols and for improving our comprehension of any tradeoffs between sulfate and nitrate control relevant to competing particle or ozone formation pathways.

To begin collecting this needed information, MANE-VU implemented the Rural Aerosol Intensive Network (RAIN), a network of enhanced monitoring stations designed to provide continuous data on the concentration, composition, and visibility impacts of fine particles. NESCAUM has coordinated the deployment of the network and this initial analysis of early RAIN data to provide states, EPA, Federal Land Managers and other interested stakeholders with an improved understanding of how RAIN fits into a longterm observing program to track and improve our understanding of visibility issues.

Among the key findings of this analysis are the following;

- 1) sulfate data has validated RAIN siting criteria and has highlighted possible issues with manufacturing quality control in the SO₄ instruments,
- 2) IMPROVE organic carbon data indicates that peak daily concentrations vary over a large seasonal range throughout the Northeast with the largest increase occurring between April and July,
- daily peak organic carbon values at the Maine site correlate moderately well with mean ambient temperature but poorly correlate with the wood smoke tracer KNON,
- 4) highly time-resolved organic carbon measurement data from both Connecticut and Maine sites show strong linkages to ozone during moderately long episodes (i.e., up to several days),
- 5) data from the Maine PAMS network supports a statistically significant link between 2-hour organic carbon concentrations and both ozone and isoprene,
- 6) organic carbon concentrations during periods of high ozone suggest that the atmospheric processes that produce ozone also help generate organic aerosols,
- 7) RAIN can reconstruct about three quarters of the ambient fine particulate mass, and site specific "scaling" factors would yield highly accurate fine mass estimates on a daily basis,
- 8) a simple two component model (organic carbon and SO₄) can successfully account for more than 90% of measured (nephelometer) visibility degradation with very good correlation on a 2-hour time-scale.

RAIN currently supports haze (and health) programs by supplying insights into co-pollutant linkages and the potential for improved air quality forecasts. Evolution and growth of RAIN will supply data that, in conjunction with emerging real-time data analysis tools and techniques, will improve conceptual models and assist in validating and generating haze control strategies.

1. INTRODUCTION TO THE RURAL AEROSOL INTENSIVE NETWORK (RAIN)

The 1999 Regional Haze Rule requires states and tribes to submit regional haze State Implementation Plans (SIPs) by 2008 to improve impaired visibility in our national parks and wilderness areas. These submissions must include the following:

- 1) an inventory of air emission sources within a state that may contribute to visibility impairment,
- 2) modeling results and data analyses demonstrating the contributions to current visibility impairment,
- 3) a long-term emissions management strategy that will improve visibility on the worst days at Federally designated Class I areas (including seven national parks and wilderness areas in the MANE-VU region), and
- 4) the presentation of existing air monitoring data as well as the development of a long-term monitoring strategy.

This technical memo focuses on the air monitoring aspect of regional haze planning. The Interagency Monitoring of Protected Visual Environments (IMPROVE) network can provide the core data for the Regional Haze program in both establishing baseline visibility conditions and in providing critical information regarding progress toward visibility goals.

We need, however, more information to successfully achieve visibility planning goals. Data will be critical for understanding the complex issues associated with the organic carbon portion of fine aerosols and add knowledge of any tradeoffs between sulfate and nitrate control relevant to competing particle formation pathways. In order to collect the needed information, MANE-VU has developed the Rural Aerosol Intensive Network (RAIN), a network of enhanced monitoring stations capable of providing continuous data on the concentration, composition, and visibility impacts of fine particles. The RAIN, deployed in 2004, will play a prominent role in future visibility control programs and the monitoring strategy requirements of Regional Haze SIPs.

RAIN is a cooperative effort of the MANE-VU member state air agencies. NESCAUM has coordinated the deployment of the network and the initial data analysis of the early RAIN data to provide states, EPA, Federal Land Managers and other interested stakeholders with an improved understanding of how the RAIN sites fit into a long-term observing program to track and improve our understanding of visibility issues.

The RAIN sites collect highly time resolved (1-2 hour) aerosol mass, composition, meteorological and optical property measurements to provide enhanced insight into the source characteristics and formation processes associated with regional aerosol. MANE-VU chose the RAIN sites for their remoteness relative to local sources, their elevation (which provides insight into atmospheric physics) and their ability to represent specific geographical areas.

By choosing moderate elevation sites, MANE-VU expects the RAIN sites will measure both surface emissions from (relatively) nearby sources as well as material aloft that had been transported long distances. For example, we anticipate that the Piney Run site (western Maryland) will provide data on the influence of large sulfur dioxide sources in the Ohio River Valley while a wide variety of sources from several locations will likely affect the Mohawk Mountain site (northwestern Connecticut). The McFarland Hill site (coastal Maine) is near the downwind edge of the continental US, thus we expect it to report on well-aged, aerosol-laden air masses.

In addition to siting considerations, the highly time-resolved (1-2 hour) RAIN data will provide good insight into the effects of various sources as well as physical and chemical processes in the atmosphere. Daily (24-hour) sampling loses much of this information.

The McFarland Hill site in Acadia National Park (Bar Harbor, ME) is situated at 44.3769 degrees north latitude and 68.2608 degrees west longitude at an elevation of 150 meters above sea level (asl). The State of Maine originally established this site in late 1997. The Piney Run site at the Frostburg Reservoir (Frostburg, MD) is at 781 m. asl (similar to surrounding terrain) and is located at 39.7060 degrees north latitude and 79.0120 west longitude. This site was established in 2004. The Mohawk Mountain site (Cornwall, CT) is at 513 m. asl and is located at 41.8331 degrees north latitude and 73.3033 degrees west longitude. It was established in the mid-1980s as part of Connecticut's PM10 network.

2. COMPARATIVE DATA ANALYSIS

Four analytical approaches are provided in this report; 1) analysis of SO₄ data, 2) analysis of carbon data, 3) reconstruction of fine mass from carbon and sulfate data, and 4) reconstruction of visibility from carbon and sulfate data. Although spatial and temporal variations exist, SO₄ is generally the largest single component of fine particle mass throughout the MANE-VU region and so we review the SO₄ data first. Carbon, especially organic carbon, is of great interest both because it is found in quantity and because its source attribution is less certain. In this analysis, we placed an emphasis on analyzing summertime organic carbon concentrations when biogenic carbon sources are active. Reconstruction of fine mass from chemical components has historically been done on an every-third-day basis. We examined the highly time resolved RAIN data for its ability to recreate the daily fine mass values captured by on-site continuous mass monitors. Lastly, because fine particles comprise the bulk of material responsible for regional haze, we examined the ability of RAIN data to reconstruct visual range and compared those results to the nephelometer data as used as in the IMPROVE program.

Due to the emerging nature of carbon measurements in general and real-time measurements in particular, a definition of terms is in order. We have adopted the following nomenclature for this report:

For RAIN data:

Total non-adjusted carbon,	TC = (thermal OC) + (thermal EC)
Organic carbon,	Opt OC = (TC) - (Opt EC)
Blank corrected organic carbon	$OC_{corr} = (Opt OC) - ("blank")$
Organic mass by carbon,	$OMC_s = ((Opt OC) - 0.5) \ge 1.8$
Total adjusted carbon,	$TC_{adj} = (OMC_s) + (thermal EC)$
For IMPROVE data:	
Organic carbon,	$OC_f = OC1 + OC2 + OC 3 + OC4 + OP$
Organic mass by carbon,	$OMC_I = (OC_f) \times 1.4$

2.1. RAIN Sulfate Data Analysis

Time series displays of SO₄ data from the RAIN sites (Figure 2-1 through Figure 2-3) demonstrate that the network is sampling different air regimes as planned. Although there is substantial sub-daily variability in sulfate concentrations at RAIN sites, we have aggregated the hourly data in Figure 2-1 through Figure 2-3 into daily values to give a sense of the seasonal data range without distracting detail. For example, the Piney Run (MD) site (Figure 2-1) shows; a) a larger range in measured concentrations during the summer months than during the winter, b) no evidence of a "basement" concentration and c) the highest "episode" occurring during the coldest portion of the year.



Figure 2-1: Maryland daily RAIN SO₄ data

As noted in the introduction, MANE-VU selected the Piney Run site to measure the impact of the large sulfur sources in the Ohio River Valley. The dynamic range and highly variable minimum values throughout the year in the Piney Run data are indicative of such sources.

Figure 2-2 shows the SO₄ time series for the next more northerly RAIN site, Mohawk Mountain (CT). The Mohawk site displays; a) large, short-term spikes during the summer, b) winter levels over $2 \mu g/m^3$ during extended periods of time (two days or more), and c) minimum values that are consistently lower than those in Maryland. Prevailing summer winds at Mohawk are southern and westerly and point toward midwestern sources but perhaps not the same ones influencing Piney Run because the peak Mohawk levels (at least for the summer of 2004) are generally higher than at Piney Run. It should be noted that higher peak values recorded at Mohawk may be due to: 1) additional time aloft that allows for more complete conversion of precursor compounds to sulfate; 2) the impact of additional sources between Piney Run and Mohawk, or 3) the relatively lower response of the Thermo sulfate instrument at Piney Run. Winter SO₄ concentrations at Mohawk are clearly lower than those at Piney Run but the distinct multi-day periods of moderate concentration indicate possible sources to the west and north of Mohawk since prevailing winter winds are from the northwest quadrant.



Figure 2-2: Connecticut daily RAIN SO4 data

Figure 2-3 shows time series SO₄ data from McFarland Hill (ME). Note that due to a negative baseline offset detected in the data, we added a correction of 0.2 μ g/m³ to hourly data for the third quarter of 2004. The McFarland data are characterized by; a) elevated summer-time SO₄ concentrations, b) minimum values throughout the year that approach the method's expected detection level of ~0.3 μ g/m³ (for daily values) and c) the lowest overall concentrations in the RAIN network. All of these attributes are consistent with the siting criteria used in selecting the McFarland Hill site.

One data feature common to all three sites is the regional "spike" visible in late August. This multi-day event resulted in the highest recorded value at Acadia, the 2nd highest value at Mohawk and the fourth highest (3rd highest if only summer values are considered) at Piney Run. During this event the peak values occurred on IMPROVE sample days (8/25 at Piney and 8/28 at Mohawk and Acadia) but for those visibility/health-oriented planners interested in detailing the onset, duration and maximum short-term values that accompany such episodes, the hourly RAIN data provides a rich and timely source of data not matched by IMPROVE.



Figure 2-3: Maine daily RAIN SO4 data

Although graphical presentations are valuable for overview purposes, they do not lend themselves to extraction of numerical data. Table 2-1 contains basic statistical information on daily SO₄ data from the three RAIN sites.

	Piney Run TECO SO₄	Mohawk Mtn. TECO SO₄	McFarland Hill TECO SO₄
Max. (ug/m3)	13.1	14.1	12.0
2nd high (ug/m3)	11.1	13.5	8.8
Mean (ug/m3)	3.38	2.13	1.41
Median (ug/m3)	2.90	1.37	0.95
Num. of Obs.	245	266	272

Table 2-1: RAIN daily SO4 statistics

The Connecticut site recorded the highest daily value during the period of record. This is somewhat surprising since the Maryland site is closest to the region thought to contain the largest sources of sulfur impacting the Northeast. Because the maximum value can be highly variable, we also include the second highest value from each site in Table 2-1. Mohawk also recorded the second highest value. This may be because the monitor is measuring the aerosol form of SO_4 (i.e., mostly ammonium sulfate), and virtually all ammonium sulfate is formed in the atmosphere via secondary pathways (cloud processing). Since this secondary process requires considerable time aloft, maximum SO_4 concentrations from Midwest sulfur sources may not have time to form before reaching the Maryland site, whereas the greater downwind distance to the Connecticut site increases the time for atmospheric processing. It is also possible that network start-up issues, unusual meteorology, or differences in monitor response to sulfate played a role in this result. Additional years of data collection will help clarify this situation. Both the mean and median statistics demonstrate the expected SW to NE gradient, with highest values in Maryland and lowest values in Maine.

2.2. IMPROVE SO₄ Comparison

Given the emerging nature of the SO_4 technique used in RAIN, it is logical to seek some indicator of accuracy. Although there is no standard for the measurement of aerosol SO_4 , we may obtain some degree of commonality by comparing RAIN results with the IMPROVE program, which has a long history of measuring SO_4 and other aerosol-based constituents. MANE-VU designed the RAIN network with such a comparison in mind and all sites have IMPROVE samplers. Although the Connecticut and Maine sites were operating IMPROVE samplers before the advent of RAIN, the Maryland site was not and thus has no 'historical' data to draw upon.

See Appendix A for details on the operating principles of the continuous SO_4 (and other RAIN) monitors. In addition, independent research by others (see Appendix C) has provided assurance that the Thermo Electron Corp., Model 5020 continuous sulfate monitor can accurately reproduce ambient SO_4 concentrations.

Because IMPROVE sites report data at "local" conditions and the RAIN sulfate data used here were reported in terms of "standard" conditions (25°C and 29.92 in. Hg), a unit conversion must be performed before making comparisons. We converted RAIN data to local conditions by use of the equation:¹

CONCLoc = CONCSTP * [298 / local temp. (°K)] * [local pressure (in. Hg) / 29.92]

Where: 298 is "standard temperature," in this case 25° Celsius or 298 Kelvin; 29.92 is standard pressure at sea level in units of "inches-of-mercury;" and $CONC_{STP}$ is the SO₄ value reported by the TECO monitor.

Before comparing RAIN to IMPROVE data, we had to make a decision regarding the form of IMPROVE sulfate to use. IMPROVE makes redundant measurements of key parameters such as SO_4 when possible. IMPROVE makes a direct measurement of sulfate (and other water soluble ions) from nylon filters and also records an indirect SO_4 measure from the X-ray analysis of Teflon[®] filters. In the latter case, the measured elemental sulfur (S) value is multiplied by 3 to arrive at an estimated sulfate value. As

¹ On-site sensors supplied local temperature data and we assumed local pressure to be the annual elevationadjusted value found in IMPROVE SOP 351, Table 351-2, Modified 10/9/97.

with all but <u>absolute</u> methods, both of these techniques have sampling and/or analysis problems and the high filter loading that occurs during peak SO_4 events exacerbates those problems. We decided to employ the technique (Nylon filter) that resulted directly in a sulfate value. We performed some cross checking with the elemental approach, but unless otherwise specified, we derived the IMPROVE SO_4 values from the analysis of water soluble ions from nylon filters.

We compared daily RAIN SO₄ data to IMPROVE SO₄ data for the three sites and plotted the results on a single graph for ease of comparison (Figure 2-4). The scatter plots for the Mohawk and McFarland sites had similar slopes with negligible offsets and a high R2 (0.93 or greater). The best-fit line slopes from a Least Squares Linear Regression (LSLR) analysis indicate that the RAIN SO₄ monitors at these two sites are detecting ~73% of the sulfate that the IMPROVE filters are reporting. Although this efficiency is lower than that reported during initial field tests of the TECO monitor, it is not unacceptably so and work continues on determining the cause of the difference.

Results from the Piney Run site were significantly different from the other RAIN sites. At Piney Run the slope was 0.52 with an intercept of 0.68 – a large departure from McFarland Hill and Mohawk. The R2 at Piney Run is smaller compared to the other sites, although still very high (0.90).

A closer look at the Piney Run scatter plot data found one day (8/10/04) where the TECO data was flagged for 11 of 24 hours. Although the reason for flagging the data did not unambiguously demonstrate that the data was invalid, we censored that day in an exploratory analysis. Three other possible outliers, clustered in the lower left hand portion of Figure 2-4 (square symbols), were found to occur on 12/20, 12/26 and 12/29. Clusters of non-conforming points are grounds for further analysis when they occur within a discreet time period since they may result from equipment malfunctions. These three days had no flags in the hourly TECO data nor in the IMPROVE data although "native" IMPROVE flags, i.e., set by site operators, are removed after initial processing. Therefore, it was not possible to ascertain whether any of the three December data points were invalid but given the temporal proximity of the data points, we removed them for an exploratory analysis.



Figure 2-4: Daily SO₄ data from IMPROVE and RAIN networks

Figure 2-5 shows the results of omitting these four data points from Piney Run. In this analysis the slope at Piney Run increases by 4%, the intercept decreases by $0.3 \ \mu g/m^3$, and the R2 increases to 0.97. Clearly these four points have an inordinate impact on the analysis and even though only one of them was documented as invalid, we question the integrity of all four points.

Figure 2-5: SO₄ results from IMPROVE and RAIN networks after censoring.



Even after extensive quality assurance investigation and data censoring, results from the RAIN monitor at Piney Run site still appear to be considerably different from the other sites. One explanation for a portion of this difference may relate to a sample flow problem. An audit of the Piney IMPROVE samplers during the third quarter of 2004 found that the "B" module flow (from which SO₄ concentrations are derived) was under-reported by 16%. This error would have caused over-reporting of IMPROVE sulfate concentrations (by ~16%). However, that error would only account for a portion of the observed discrepancy and the issue remains under review.

The rate of formation of sulfate is significantly temperature dependent, with the rate increasing with warmer temperatures. Because it is possible that sampling-related artifacts may also be temperature dependent, we made a comparison of IMPROVE and RAIN data on a seasonal basis.

Figure 2-6 compares the scatter plot of SO_4 from IMPROVE (Ionic) vs. RAIN for the warm season (Jul.-Sep.) versus the similar plot for the cool season (Oct. – Dec.). Figure 2-6 plots data for only one site because all three sites showed the same pattern with a higher slope for the cool season. Analysis of 95% confidence levels (not shown) of mean cool/warm pairs showed no systematic impact of temperature. The only instance where 95% confidence levels for mean IMPROVE and RAIN SO₄ did not overlap was in the warm season at the Maryland site. That result, indicating a significant difference in means only during warm periods, is difficult to interpret even in light of the IMPROVE sampler flow error discussed above.



Figure 2-6: Seasonal SO₄ comparison of Maine data.

Some measurement techniques may overload or saturate at high concentrations. A case in point occurred when the IMPROVE program, which had consisted solely of rural sites, began operating in urban/suburban areas. The elevated aerosol concentrations in urban areas overloaded the IMPROVE nylon filters. IMPROVE subsequently increased the filter size to accommodate the higher loadings. In line with that knowledge, an analysis was undertaken to examine the performance of the IMPROVE/RAIN networks at high SO₄ concentrations.

To assess the impact of high sulfate concentrations on IMPROVE/RAIN method performance, sample pairs at each site (6/29 - 12/29/04) were sorted by descending concentration of the IMPROVE (ionic) sample. The 10 highest samples at each site were labeled "High" and the remaining samples (40-51 depending on site) were considered "Low." We calculated the 95% confidence levels (Table 2-2) for the mean SO₄ of the high and low days at each site for both the IMPROVE (ionic) and RAIN (TECO) methods. The only instance where the upper and lower 95% confidence levels did not overlap was for "high" days at Piney Run. This indicates that the difference between IMPROVE and RAIN mean values on "high" days at Piney is statistically significant.

One possible explanation is that there are some SO_4 losses when the aerosol is strongly acidic (approaching sulfuric acid). In general, that is the case when sulfate levels are high. There is limited evidence at the Piney site that the Model 5020 sulfate response, compared to filter sulfate, is somewhat suppressed at higher levels, but no clear conclusions can be made from the data to-date. Only very limited and early tests with sulfuric acid aerosol have been done on the Thermo sulfate method (showing some reduced response), so this remains an open question.

Table 2-2: 95% confidence levels for mean IMPROVE and RAIN SO4 values ondays with either high or low concentrations.

	Piney Run IMP SO₄ (ionic)	Piney Run RAIN SO₄	Mohawk IMP SO₄ (ionic)	Mohawk RAIN SO₄	McFarland Hill IMP SO4 (ionic)	McFarland Hill RAIN SO₄
High Day Mean +95% CL	13.7	7.9	12.7	9.5	8.1	6.2
High Day Mean -95% CL	10.4	6.1	6.3	5.1	3.4	2.6
Low Day Mean +95% CL	4.3	2.8	1.9	1.5	1.1	1.1
Low Day Mean -95% CL	2.7	1.9	1.4	1.1	0.78	0.73

2.3. Conclusions

From the foregoing analyses we conclude that;

- daily mean SO₄ concentrations measured at RAIN sites during 2004 decreased along a Southwest-to-Northeast transect supporting a prime network siting criteria regarding source attribution,
- 2) the decrease in daily <u>minimum</u> values along the transect was especially evident in time series plots further validating siting criteria,

- reported sulfate concentrations from the RAIN monitors, when compared to IMPROVE direct sulfate measurements (water soluble ions), agreed to within a mean of 72% at two sites and 56% at the third (Maryland) site,
- sulfate results from the RAIN monitors at the Maine and Connecticut sites appear linear across the range of concentrations encountered and insensitive to seasonal effects,
- 5) the Maryland IMPROVE sampler may have experienced reduced response to sulfate aerosols that are very acidic [e.g., approaching sulfuric acid],
- 6) RAIN sulfate data are highly correlated with IMPROVE sulfate data at all sites, allowing for linear data correction techniques to make the RAIN sulfate data consistent numerically with the IMPROVE sulfate data,
- 7) RAIN provides highly time-resolved SO₄ data for the Northeast region that can assist SIP planners involved in Regional Haze control.

The observation that both ozone and aerosol-borne organic carbon follow a similar diurnal pattern at RAIN sites (see Figure 3-1) prompted further analysis.

Figure 3-1: Diurnal (2-hour) patterns at Mohawk Mtn., CT



This observed correlation was especially strong during "episodes" as indicated in Figure 3-2 by both the time series and the summary of the least-squares-linear-regression (LSLR) analysis (upper right hand corner). Note that; 1) there does not appear to be a significant lag or lead of either component and 2) the statistical significance of this correlation over the entire data set currently available (7/22-9/31/04) is not robust. The LSLR correlation (expressed as R²) for the entire 7/22-9/31/04 period is less than half (0.36) of the "episode" value. Analysis of longer O₃/OC_{corr} time series (not shown) suggests that downward "spikes" in OC_{corr} during periods of relatively low carbon concentration (i.e., less than 4-5 µg C/m³) may be driving the poor correlation.



Figure 3-2: Example of episodic ozone vs. organic carbon – Mohawk Mtn. CT

We examined organic carbon data (OC_{corr}) from the RAIN site at McFarland Hill in some detail due to the availability of; 1) historical IMPROVE carbon data from that site and 2) Photochemical Assessment Monitoring Station (PAMS) data from a nearby site also in Acadia NP. Although the RAIN and IMPROVE networks use different sampling and analysis techniques for organic carbon, it can be argued that the IMPROVE organic carbon data (OC_f) resembles the RAIN OC_{corr} data. Although shorter in historical length, we viewed the highly time-resolved (hourly) PAMS volatile organic carbon (VOC) data set as a tool that might provide insight into the difference in contributions from man-made and natural sources.

As a first step in investigating the possible link between O_3 and OC_f at the local level, IMPROVE data was screened to remove instances of known or suspected bias from short-term, remote sources such as wildfires. This censoring approach employed both a high OC_f limit as well as the ratio of OC_f -to-fine mass to identify biased samples. We considered samples with OC_f values greater than or equal to 5.0 µg/m³ as "high" and if the OC_f -to-fine mass ratio for high samples exceeded 0.333, we considered those samples to be biased. This screening process resulted in removing six data points from the original 1,763 sample set at Acadia that spanned the period 3/2/88 to 12/29/04. The six censored dates are 8/9/89, 8/14/91, 8/23/95, 6/11/97, 6/14/97 and 7/10/02.

We applied the same censoring criteria to data from other sites in the Northeast region subsequently added to the analysis. OC_f values for samples collected on 11/12/01 and 11/18/01 from the Great Smoky Mountains NP (GRSM) site met the censoring criteria and the 11/15/01 sample from that site had a "high" OC_f value but the fine mass concentration was not available. Those three samples were removed from the GRSM data set which originally consisted of 1,801 samples that spanned the same time period as Acadia.

3.1. IMPROVE Seasonality

An initial look at mean monthly IMPROVE OC_f data from Acadia showed extreme seasonality with summer (JJA) levels often exceeding the sum of the remaining nine months (Figure 3-3). More often than not, the hot and cold seasons tracked each other directionally (Figure 3-4) but the summer/winter ratio (not shown) was highly variable over the period of record, indicating that a long-term (annual or longer) driver was not present.

Organic carbon concentrations measured by IMPROVE at the Acadia site were observed to increase sharply between the months of April and July (Figure 3-4). Typically, for the period of record, the mean March-May OC_f is the lowest three-month average for that calendar year while the three months immediately following (June-August) have the highest three-month average. We hypothesize that variation in meteorological parameters (e.g., rainfall, wind speed, wind direction) are responsible for this large gradient. It will require a detailed examination of meteorological data from Acadia to validate this hypothesis, which has yet to be undertaken.



Figure 3-3: IMPROVE organic carbon data from Acadia, ME.



Figure 3-4: Mean seasonal behavior of OC_f at Acadia, ME.

3.2. Temperature Dependence

Temperature is a potential driver for increased summer OCf for both chemical and biological reasons; biogenic VOC sources and the photochemical processes to convert VOC to secondary organic carbon aerosols both peak in the warm summer months. We compared mean and maximum monthly temperatures to monthly mean OC_f concentrations from McFarland Hill. Although some inter-annual variation exists, neither the mean monthly temperature nor the maximum monthly temperature account for the year-to-year variation in summer OC_f peak values at Acadia.

The scatter graph of maximum monthly temperature versus mean monthly OC_f for the period March 1988 through August 2004 (Figure 3-5) shows a large degree of scatter especially during the warmest periods. Further investigation involved dividing annual data into "cool" and "warm" periods of various lengths and overlaying the resulting scatter graphs. Figure 3-6 shows the maximum temperature vs. OC_f scatter for the six warmest months (May, June, July, August, September, and October) overlaid on the scatter for the six coolest months. The pattern for cool months appears to be distinctly different than that for warm months.



Figure 3-5: Maximum monthly temperature vs. mean monthly OC_f correlation.



Figure 3-6: Cool vs. warm temperature impact on OC_f – Acadia.

We tried other monthly combinations and somewhat surprisingly, the best correlations did not occur during the extreme periods. For example, when we compared the four hottest months (June, July, August, and September) to the other eight months, the correlation in the cooler months improved slightly but the "hot" correlation became worse. We found the optimal correlation was when the "warm" season consisted of April, May, June and July (Figure 3-7).



Figure 3-7: Optimal temperature vs. OCf distribution.

Although OC_f inversely correlates with the coldest mean monthly temperatures and directly correlates with the warmest months at the McFarland Hill site, it is not clear why the hottest months do not yield the best correlations in this analysis.

3.3. Wood Smoke Impact

In the above analyses, the focus was on local conditions thought to influence the formation of aerosol-borne organic carbon. To screen out the impact of distant events such as wildfires, we censored some samples from those analyses. However, wildfires (and to a lesser extent prescribed burning) occur throughout the Northeast. We do not know the impact of those occurrences or whether they exceed the impact of larger, more distant wildfires. Therefore, we examined the IMPROVE "non-soil potassium" parameter known as KNON in conjunction with OC_f data.

KNON is thought to be a reasonable tracer for wood smoke, but has significant limitations. It is calculated as the potassium concentration minus 60% of the iron concentration [KNON = K – (Fe*0.6)]. We calculated monthly mean KNON values for the McFarland Hill site and compared them to monthly OC_f concentrations (Figure 3-8). For this comparison, we did not censor any values from the OC_f mean concentrations since the impact of smoke was of interest. We see some of the largest KNON peaks coinciding with the greatest OC_f peaks but KNON also possesses a large cold weather signal – possibly due to local wood combustion. The cold weather KNON pattern is in part responsible for the poor KNON/ OC_f correlation (see LSLR analysis data in upper right portion of Figure 3-8). However, even when we removed the eight coldest months from the correlation (not shown), results improved only marginally. These results suggest that in the long term, wood smoke does not significantly influence summer OCf concentrations at Acadia based on KNON analysis.



Figure 3-8: Mean monthly KNON vs. OC_f concentrations – Acadia.

However, when the monthly patterns of EC from IMPROVE at Acadia and two other rural Northeast US sites is considered (Figure 3-9), there is evidence of a significant summer increase in EC also, with the seasonal pattern of EC at Acadia similar to that of OC (Moosehorn has some winter EC peaks, perhaps from local wood-burning). Since biogenic sources do not generate EC and there is no space heating during the summer, that leaves either mobile sources (a source of EC and OC) or woodsmoke from wildfires (a source of both EC and OC, but dominated by OC). For Acadia, either of these sources would have to be long-distance transport. Wildfire season in North America is summertime, so very-long range transport of wildfire smoke is a plausible source of at least some of the observed summertime increase in rural OC. Another potential source of summer woodsmoke at Acadia could be from local campfires, but since the EC and OC data from the Lye Brook IMPROVE site show similar seasonal patterns (see Figure 3.9 below and the discussion in section 3.4), any influence of local woodsmoke at the Acadia sampling location should be considered relatively minimal in the context of this analysis.

Figure 3-10 and Figure 3-11 show the likely source region associated with EC and OC during June-July-August based on a CATT analysis of IMPROVE data from Acadia². The similar pattern coming from the Northwest for both pollutants suggests that a common woodsmoke source is likely to be, in part, responsible for their prevalence during summer months, despite the poor KNON correlation. The same analysis for sulfate and sodium is also available.³

² http://datafedwiki.wustl.edu/index.php/CATT_IP

³ http://datafedwiki.wustl.edu/index.php/CATT_ACAD_SummerOCEC



Figure 3-9: Mean monthly EC concentrations – Acadia, Lye Brook, and Moosehorn IMPROVE sites (1995-2004).

Figure 3-10: Potential Source Region for EC at Acadia National Park between 1995 and 2005 during June, July August (Courtesy of CATT)



Figure 3-11: Potential Source Region for OC at Acadia National Park between 1995 and 2005 during June, July August (Courtesy of CATT)



3.4. Regional Organic Carbon

To determine whether the seasonal pattern of organic carbon at Acadia was anomalous, we examined OC_f data from five other IMPROVE sites in the eastern U.S.. Those additional sites are Brigantine, NJ (BRIG); Dolly Sods, West Virginia (DOSO); Great Smokey Mtn., TN (GRSM); Lye Brook, VT (LYBR) and Shenandoah, VA (SHEN).

We calculated running three-month means to filter out some of the short-term noise and to indicate whether the sites have similar seasonal patterns and/or similar dynamic data ranges (Figure 3-12). Because extreme values are so variable, we calculated a "dynamic range" for each site by finding the 10th and 90th percentiles for the monthly means and then taking the ratio of the 90th/10th percentiles. This ratio shows that the Acadia site is near average (2.74) for the region while SHEN (2.76) and LYBR (3.01) exceed ACAD's dynamic range and DOSO (1.87), BRIG (2.15) and GRSM (2.42) fall short of it.

Not surprisingly, ACAD correlates best with the two closest sites, LYBR ($R^2=0.58$) and BRIG ($R^2=0.44$). However, the LYBR/SHEN site-pair (not shown) displays a correlation about equal to ACAD/LYBR and possesses a "better" slope (0.99).



Figure 3-12: Historical regional organic carbon concentrations.

As one might expect, there is a fairly large variation in the dynamic range of values at the sites examined. However, the strong seasonal variation noted at the Acadia site is not so different from other sites to invalidate the analyses performed in sections 3.1-3.3. It appears therefore, that the processes driving the large spring-to-summer OC_f variations at Acadia are also at work at other IMPROVE sites in the Northeastern U.S.

3.5. Semi-Continuous Organic Carbon vs. Ozone

Because of the low aerosol carbon loadings at the rural RAIN sites, the RAIN program configured the Sunset carbon analyzers to collect two-hour samples, minimizing the impact of the OC blank value. We aggregated hourly ozone values from the McFarland Hill site into equivalent two-hour averages to facilitate comparison. The limited real-time organic carbon data (OC_{corr}) collected to-date precluded seasonal analysis so we performed daily (day-of-week) and additional diurnal (hour-of-day) analyses.

To examine the possible influence of daily local source impact (i.e., mobile source traffic) on the site, we generated a day-of-week graph (not shown) of O_3 vs. OC_{corr} . Although a large difference between OC_{corr} and O_3 was evident on Thursdays,

that difference was likely due to missing data because there were 100 "Thursday" samples while the number of observations on other days ranged from 112 to 141.

We generated a plot of mean OC_{corr} by day-of-week (not shown) that displayed some weekday vs. weekend differences, especially during the overnight hours. However, analysis of mean values at 95% confidence levels (not shown) found the differences to be insignificant indicating that the site was relatively free from local source impacts.

If the photochemical processes that create O_3 are also associated with the formation of aerosol-based organic carbon, then that linkage should be evident during periods of high O_3 . To check this hypothesis, we classified days as "high" or "low" based on the maximum hourly ozone concentration (i.e., the break point) observed that day. A LSLR analysis comparing O_3 and OC_{corr} on "high" and "low" days demonstrated that high ozone days always showed a stronger O_3/OC_{corr} correlation than low days (regardless of the 'break point' chosen). We did not find, however, that those correlations were statistically significant even on the 97th percentile days (Figure 3-13and Figure 3-14) when the low number of data points became the controlling factor.



Figure 3-13: McFarland Hill ozone vs. OC_{corr} for days with 60ppb (O₃) breakpoint.



Figure 3-14: McFarland Hill ozone vs. OC_{corr} for days with 69ppb (O₃) breakpoint.

3.6. PAMS Data Comparison

We examined ozone-relevant data from the nearby PAMS site on Cadillac Mt. in Acadia NP (~2 miles away and 1000 feet higher in elevation) in conjunction with the McFarland Hill organic carbon data. Although PAMS data from Acadia typically encompasses only the summer months (June-September), in more recent years it also included the month of April. Because of its high time resolution (hourly), PAMS data is especially useful for comparison to the 2-hour OC_{corr} data captured by RAIN.

We used PAMS data for the period 6/97-9/04 in this analysis. We selected target compounds based on the likelihood of contributing to the formation of aerosol organic carbon (Song et al. 2005, Lim et al. 2005) and included isoprene, toluene, m/p xylene, aromatics, total organics (TNMOC), NOx and the ratio of TNMOC/NOx.

We arranged data on the compounds of interest into hour-of-the-day bins to create diurnal patterns. We used only days with at least 75% data capture (i.e., 18 or more valid hours). Figure 3-15 and Figure 3-16 show diurnal patterns for seven PAMS-related targets. An intrinsic value in diurnal plots lies in the patterns produced. It is not necessary for compounds arising from the same (or similar) processes to exhibit peaks and minima at exactly the same hour of the day since atmospheric reactions proceed slowly, rather it is the shape of the daily distributions that point toward relations.



Figure 3-15: Diurnal pattern of low concentration PAMS compounds.



Figure 3-16: Diurnal pattern of high concentration PAMS compounds.

Visually, the PAMS targets that most closely followed diurnal OC_{corr} pattern at McFarland Hill were isoprene followed by ozone (see Figure 3-17). Because we previously examined above the relationship between mean daily ozone and organic carbon from the IMPROVE program (OC_{f}), this analysis focused on a possible link between RAIN OC_{corr} and isoprene.



Figure 3-17: Diurnal O₃ and OC_{corr} from McFarland Hill vs. Isoprene from Acadia.

We did not attempt a direct correlation between isoprene and two-hour OC_{corr} measurements because different sites measured those parameters. Instead, we defined "high" isoprene days by the highest daily one-hour isoprene concentration at Cadillac and then matched those "dates" with OC_{corr} data from the nearby RAIN site. We then divided dates into "High Isoprene" and "Low Isoprene" categories based on their percentile rank. At all of the percentiles examined (50th, 75th, and 95th), the mean OC_{corr} concentration at McFarland Hill, on days when the Cadillac site exhibited high isoprene values, was significantly higher (at the 95% confidence level) than on days when Cadillac isoprene was low (Figure 3-18). This suggests that the process(s) that drive(s) isoprene production also participate(s) in OC_{corr} formation.

Finally, we compiled all available two-hour average isoprene and OC_{corr} data for the period 7/6/04 - 9/30/04 and arranged them sequentially. The LSLR equation for those 12 data pairs resulted in a line with a slope of 0.24, a Y-intercept of 1.23 and an R² of 0.70. This relatively high correlation suggests that the processes responsible for isoprene production are linked to organic carbon formation at Acadia.



Figure 3-18: Mean daily organic carbon from McFarland Hill (with 95% confidence levels) on days with high and low isoprene (measured at Cadillac Mtn.).

3.7. Conclusions

Analysis of RAIN organic carbon data in conjunction with comparison to IMPROVE organic carbon data indicates that;

- 1) peak daily organic carbon concentrations have a very large seasonal range at IMPROVE sites throughout the Northeast,
- 2) the largest annual increase in organic carbon occurs between April and July,
- 3) at the Acadia site, local (mobile) sources do not significantly influence OCf data, peak values correlate moderately well with mean ambient temperature but poorly correlate with the wood smoke tracer KNON,
- 4) at Acadia, IMPROVE EC shows a summer increase similar to OCf; this could be long-range transport from either mobile or wildfire sources, but implies that biogenic sources are not the sole contributor to the observed summer OCf observed at the Acadia site,
- 5) highly time-resolved organic carbon measurement (OC_{corr}) data from both Mohawk and Acadia show strong linkages with ozone during moderately long "episodes" (i.e., up to several days),

- 6) OC_{corr} concentrations from the RAIN site at Acadia correlate with both ozone and isoprene at statistically significant levels,
- 7) the elevated organic carbon concentrations recorded by both IMPROVE and RAIN during periods of high ozone suggest that the atmospheric processes that produce ozone also help generate organic aerosols,
- 8) the above-noted links need to be further investigated so that air quality planners can craft fine aerosol control strategies without concern over inadvertently increasing ozone (or other) pollutants.

4. RECONSTRUCTION OF FINE MASS

Differences in the type of continuous fine mass monitors employed at the RAIN sites as well as the amount of historical $PM_{2.5}$ (FRM) data available resulted in slightly different data analysis approaches at each site. The McFarland Hill site at Acadia is equipped with a TEOM operated at 50°C and has fine particle FRM data that precede start-up of the RAIN SO₄ monitor. The Mohawk and Piney Run sites both use Beta Attenuation Monitors (BAM) to monitor fine particle mass. The FRM $PM_{2.5}$ sampler at Mohawk was brought on-line on 1/1/05 but the FRM at Piney did not begin operation until the second quarter of 2005. The chief difference in reconstructing fine mass arises from the availability (or lack) of "FRM-like" continuous measurements. McFarland Hill has such FRM-like continuous measurements for the entire RAIN program but the Mohawk and Piney sites do not.

4.1. McFarland Hill, Maine

Continuous RAIN sulfate data are available from McFarland Hill for 6/26/04 - 3/31/05 while continuous carbon data exist for the 7/7/04 - 3/31/05 timeframe. Due to missing or invalid data, reconstructed fine mass data are available for only 203 out of a total 267 days (76%) during the 7/7/04 - 3/31/05 analysis period.

Both the IMPROVE and RAIN programs measure organic carbon and sulfate but RAIN has no capacity to assess either the "soil" (SOIL) or "ammonium nitrate" (ammNit) component of fine mass. As part of the fine mass reconstruction effort, we compared the RAIN-derived reconstructed fine mass to IMPROVE fine mass. In order to make that comparison as accurate as possible, we compiled monthly average values for ammNit and SOIL from historical IMPROVE data (~ 17 years for the McFarland site).

To reconstruct fine mass, we multiplied the RAIN SO₄ concentrations by 1.37 to convert sulfate mass to the equivalent mass of ammonium sulfate. (As with the IMPROVE program, we assumed that all measured sulfate was in the form of ammonium sulfate). We multiplied RAIN organic carbon concentrations (OC_{corr}) by 1.8 to convert carbon mass to volatile organic carbon mass (OMC_s). This is a departure from IMPROVE, which has historically used a carbon factor of 1.4 but we felt that emerging information (Hand and Malm, 2005) supported the higher factor for eastern rural sites. Lastly, we added the elemental carbon fraction (thermal EC) to the OMC_s fraction to create a "total adjusted carbon" value (TC_{adj}).

In the final step, we added the two RAIN components (ammSO₄ and TC_{adj}) to the IMPROVE "SOIL" and "ammNit" mean monthly values. We then compared the resulting reconstructed daily fine mass values to the FRM-like TEOM (McFarland Hill) concentrations (or BAM concentrations at Mohawk and Piney). We also compared the reconstructed RAIN mass with the every-third-day IMPROVE fine mass values.

Figure 4-1 shows the results of the TEOM-to-reconstructed mass exercise. Note that this simple approximation technique captures nearly 80% of the daily fine mass. The R2 is also extremely robust (0.86), indicating high confidence in the statistical significance of this approach. Although a small number of points in the upper range can bias the LSLR approach, this does not appear to be the case here since the highest concentration points are scattered equally above and below the best-fit line.

Although not shown, the LSLR analysis of IMPROVE fine mass (from Teflon[®] filters as opposed to the TEOM) against reconstructed RAIN mass yielded results nearly identical to those above. The best fit line slope in this case was 0.79 with an intercept of +1.13 and an R² of 0.88.







Figure 4-2: Reconstructed fine mass from IMPROVE versus RAIN reconstructed daily fine mass at McFarland Hill.

We also compared the IMPROVE reconstructed fine mass to the reconstructed RAIN equivalent. Figure 4-2 shows the scatter plot of that comparison as well as the results from the LSLR analysis. In this case, the R^2 is nearly identical to the TEOM comparison above but the slope has increased to 0.90, indicating that RAIN is successfully accounting for 90% of the IMPROVE reconstructed fine mass.

These analyses demonstrate a high degree of agreement between the RAIN and IMPROVE programs with respect to reconstruction of fine mass. The mass "missing" from the RAIN mass reconstruction is almost certainly due to the fact that the RAIN sulfate monitor only reported 74% of the SO₄ found by IMPROVE at the McFarland site (see SO₄ analysis section above).

4.2. Mohawk Mountain, Connecticut

The scatter plot in Figure 4-3 shows BAM fine mass versus reconstructed mass for the Mohawk site. The simple reconstruction model does not work as well here, predicting just under 60% of the fine mass but the R2 is still robust. Note also that BAM monitors tend to produce higher fine mass concentrations when co-located with FRM samplers while TEOMs record lower values especially during cold weather. These different continuous PM2.5 measurement technologies at Mohawk and McFarland contribute to this between-site discrepancy in fine mass reconstruction.



Figure 4-3: Mean daily BAM fine mass at Mohawk Mtn., CT, versus reconstructed fine mass from the RAIN program.

4.3. Piney Run, Maryland

The scatter plot in Figure 4-4 shows BAM fine mass versus reconstructed mass at the Piney Run site. Due to the limited IMPROVE data available at the Piney site, we could not calculate meaningful "SOIL" and "ammNit" factors. For the Piney Run analysis, we derived those factors from the nearby Dolly Sods Wilderness IMPROVE site in West Virginia that has been operating since 1991. As with the Mohawk site, the reconstruction algorithm does not work as well as at McFarland Hill. However, the Piney Run site continuous mass monitor is also a BAM and the cautionary note regarding measurement method differences at Mohawk also applies here. Another factor is the relatively reduced response to sulfate from the Piney Run Thermo sulfate monitor. One other point regarding the Piney Run data involves censored data points. As noted in the SO₄ section above, it appeared that there were three "outliers" in the daily sulfate data. For this analysis, we excluded two of those daily values (based on their status as SO₄ outliers). The third date did not have valid carbon data and we therefore did not include it in the reconstruction data set.



Figure 4-4: Mean daily BAM fine mass at Piney Run, MD, versus reconstructed fine mass from the RAIN program.

One way to investigate the above-noted differences due to continuous mass measurement techniques is to utilize a measurement method (not necessarily continuous) that is common to all sites. Because all RAIN sites have IMPROVE samplers that measure fine mass (~PM_{2.5}), we undertook a comparison of reconstructed RAIN mass with IMPROVE fine mass to assess the impact of the different continuous techniques.

Because IMPROVE operates on an every-third-day schedule, the number of data points available for comparison is reduced. The lag in validating IMPROVE data is also a factor – only data through the end of calendar year 2004 was available for this analysis. There are 38-48 IMPROVE samples (depending on site) available for comparison in the 7/2/04 to12/29/04 period.

Figure 4-5 shows the results of the LSLR analyses of IMPROVE fine mass and the fine mass reconstructed from RAIN data. The range of slopes for in this case is much tighter than the range resulting from use of the on-site continuous fine mass RAIN monitors. Using a uniform fine mass measurement technique significantly reduces the variation in this simple model performance (as measured by LSLR line slope) from $\pm 16\%$ for continuous monitors to $\pm 7\%$ for IMPROVE samplers. This should not be interpreted

to mean that the IMPROVE technique is superior, rather it is a statement in favor of a <u>uniform</u> measurement technique.





4.4. Conclusions

Comparison of RAIN reconstructed fine mass data indicates that;

- when using the on-site continuous fine mass monitor as the "baseline", the Mohawk and Piney Run sites correctly identify ~60% of the fine mass while the McFarland site reconstructs 79%,
- 2) when using the every-third-day IMPROVE fine mass as the baseline metric the Mohawk and Piney sites report 65-70% of the IMPROVE mass while the McFarland results are unchanged,
- 3) although RAIN near real-time measurements somewhat under-report fine mass concentrations, the high correlations indicate that site specific "scaling" factors can be employed to yield highly accurate fine mass estimates on a daily basis,
- 4) accurate daily fine mass estimates delivered in near real time can help protect public health by improving the accuracy of air quality forecasts.

5. VISIBILITY RECONSTRUCTION

Air quality planners implemented both the RAIN and IMPROVE networks to develop a more thorough understanding of the sources and mechanisms driving regional haze. The primary metric in the regional haze program is visual range and it is important to understand how well the emerging real-time techniques employed by RAIN can reconstruct visibility impairment. The following section describes how we used RAIN data to reconstruct visibility at the McFarland Hill site.

5.1. Approach and Discussion

In addition to carbon and sulfate data, visibility reconstruction requires two other types of information; 1) relative humidity (RH) data and 2) light scattering data. A Rotronic sensor (Model MP-101A, with active aspiration) collected relative humidity (RH) data on-site. An IMPROVE nephelometer (Optec NGN-2a, operated by the National Park Service) obtained light scattering data. All three RAIN sites measure local RH and light scattering, but McFarland Hill is the only site located within a Class 1 area as defined by EPA; therefore nephelometer data from Acadia was fast-tracked for this analysis.

Although gas-phase compounds participate to some degree in the impairment of visibility, aerosols are by far the most important factor. In order to better understand the relative contributions of the major aerosol components, we evaluated sulfate and carbon data separately before examining their combined effect on visibility.

The Thermo ("TECO") Model 5020 monitor reports only "sulfate" mass (at standard conditions). For visibility purposes, the reported sulfate mass must be converted to an equivalent mass of ammonium sulfate, $(NH_4)_2SO_4$. This is done because, as previously noted, the IMPROVE program assumes that all sulfate is in the form of ammonium sulfate. In the rural Northeast during high sulfate events, this is not always true. It is, however, a reasonable first assumption because aerosol is rarely more acidic than ammonium bisulfate (NH₄HSO₄), which overall is similar to ammonium sulfate for scattering reconstruction purposes.

Additionally, $(NH_4)_2SO_4$ readily absorbs and emits water as a function of atmospheric humidity. This affects the size of the particle, which in turns affects the aerosol's light scattering ability; therefore knowledge of relative humidity is important in visibility reconstruction. Given the highly non-linear relation between relative humidity and ammonium sulfate particle size, we excluded RHs equal to or greater than 95% from this analysis. Of the 6,698 hourly observations possible from 7/1/04 through 3/31/05, this RH exclusion coupled with nephelometer data "flagged" as invalid for other reasons, amounted to 2434 hours. This left 4,246 observations, which is equivalent to a data capture rate of 63%. This is substantial given the meteorological conditions prevailing at the RAIN sites.

To begin the analysis, we multiplied sulfate concentrations by 1.37 to convert them to a mass equivalent of ammonium sulfate. This new variable (SULFATE) is essential because of the extreme size-dependent nature of $(NH_4)_2SO_4$ light scattering,

which in turn, is highly (and non-linearly) related to atmospheric RH. Next, we converted the hourly RH values to a relative humidity function f(RH) by using the conversion table adopted by IMPROVE for this purpose

(see: http://vista.cira.colostate.edu/improve/Tools/humidity_correction.htm).

Next, the standard IMPROVE dry specific scattering coefficient of 3

(see: http://vista.cira.colostate.edu/improve/Tools/ReconBext/reconBext.htm)

was applied to the hourly SULFATE values. We show below the equation used to convert sulfate mass to extinction by ammonium sulfate:

Reconstructed [$(NH_4)_2SO4$] Scattering = 3 * f(RH) * SULFATE

Note that the total number of observations available was further reduced due to lost and invalidated data from the sulfate monitor. When we compared the reconstructed visibility due to the sulfate component to the nephelometer data via LSLR, we obtained an adjusted correlation (R^2) of 0.90 along with a slope of 0.64 and a y-intercept of -1.87 (see Figure 5-1). This analysis indicates that sulfate alone (as reported by the TECO 5020) is only a fair surrogate for but an excellent predictor of visibility at McFarland Hill for this period of time.

Figure 5-1: McFarland Hill nephelometer visibility data versus reconstructed visibility from only sulfate.



Because elemental carbon absorbs rather than scatters light, it does not enter into the scattering equation although it is a factor in visibility degradation. In attempting to recreate the scattering observed by the NGN-2a nephelometer, we use only the light scattering carbon (OC_{corr}) reported by the Sunset Model 3 carbon analyzer.

The IMPROVE program uses the following equation to describe the impact of light scattering carbon:

Reconstructed Carbon Scattering = $4 * f_{org}(RH) * [OMC]$

where the dry scattering coefficient of carbon is set at 4, the relative humidity factor is set at unity (due to the weak hygroscopicity of organic carbon) and OMC is organic mass by carbon which is derived as ($OC_{corr} \times 1.8$). Note that IMPROVE has historically employed a multiplier of 1.4 (within the OMC factor) to account for the other elements associated with measured carbon mass but that factor has been subject to much debate. IMPROVE has recently (Hand and Malm, 2005) recommended increasing the carbon factor to either 1.7 or 1.8 depending on the nature (i.e., urban vs. rural). Due to the highly rural nature of the McFarland Hill site, and the aged air masses that are experienced there, we felt that a higher factor would be more appropriate, so this analysis uses 1.8 as the carbon factor.

The RAIN sites collected carbon data over a 2-hour period, so we had to perform some averaging in order to combine the carbon data with the hourly sulfate data in a reconstruction equation. We averaged the TECO SO_4 data, the NGN-2a data and the f(RH) data all over the same 2-hour periods.

In addition, an appropriate filter blank value for the Sunset Laboratory OC data was needed. Empirical data gathered to date from RAIN indicated that a factor of $0.5 \ \mu g/m^3$ was appropriate for the Model 3 in rural areas and consequently we subtracted that value from the OC data prior to use in reconstruction. Parameters from a LSLR of OMC on NGN2a 2-hour visibility data are: Slope: 0.14; intercept: 6; R2: 0.27; these results indicate that OC is not as well correlated with scattering and that the overall contribution to scattering is substantially less than that from sulfate at this site.

Figure 5-2 shows the results of the reconstructed visibility using both sulfate and carbon data.



Figure 5-2: McFarland Hill nephelometer visibility data versus reconstructed visibility from both sulfate and carbon.

Adding the carbon extinction to the sulfate extinction increases model accuracy by 14% while yielding essentially the same correlation. Note also that this model encompasses a significant portion (i.e., >8 months) of an annual cycle and thus may well represent an entire year. The combined sulfate and carbon extinction data account for nearly 80% of the visual extinction (as measured by the NGN-2 nephelometer) from the McFarland Hill site.

As noted earlier, the Teco 5020 sulfate data is lower than the IMPROVE sulfate data, but is highly correlated. A correction factor of 1.3 (based on the Teco vs. IMPROVE sulfate regressions presented earlier) makes the Teco sulfate data "IMPROVE"-like. Since sulfate dominates the light scattering at Acadia, it is worth examining the reconstructed extinction with this correction factor applied. Figure 5-3 is the same as the regression shown above, but with this correction factor included in the Teco sulfate data. The regression slope increases substantially to 0.91. This is still less than unity, as expected given the lack of other light scattering components in the modeled scattering. The R2 improves slightly to 0.94, because of the greater dynamic range of the reconstructed scattering or the relatively reduced influence of the OC scattering component (as noted above, OC is not well correlated with the 2-hour nephelometer data at this site).



Figure 5-3: McFarland Hill nephelometer visibility data versus reconstructed visibility from both sulfate and carbon, with sulfate correction applied.

5.2. Conclusions

Analysis of semi-continuous organic carbon and sulfate data with respect to measured visual range at the McFarland Hill site demonstrates that a simple two component model can successfully account for more than 90% of the visibility degradation as measured by nephelometer. The high correlation for this 2-hour data set provides increased confidence that the original IMPROVE f(RH) function is reasonable for sulfate at this site and season.

The ability to track visibility in near real time can assist analysts in better understanding atmospheric processes such as haze formation and transport. By noting episode while they are occurring analysts will be able to capture additional information that might be lost to an "after the fact" analysis. This is especially true given the emergence of new data sources and processing tools such as the DataFed project initiated by Dr. Rudolph Husar at Washington University, St. Louis (see: <u>http://datafed.net/</u>). Faster, more inclusive analyses will lead to better conceptual models and more accurate source apportionment of haze precursors.

6. REFERENCES

SOP 351 Data Processing and Validation, 10/97, (accessed 1/26/06), http://vista.cira.colostate.edu/improve/Publications/SOPs/ucdsop.asp

Hand, J. and Malm, W., May 24, 2005, *Review of the IMPROVE Equation for Estimating Ambient Light Extinction Coefficients*, (accessed 1/26/06), http://vista.cira.colostate.edu/improve/Publications/GrayLit/016_IMPROVEeqReview/I MPROVEeqReview.htm

Song, C., Na, K. and Cocker, D., *Impact of the Hydrocarbon to NOx Ratio on Secondary Organic Aerosol Formation*, Environ. Sci. Technol., 2005, 39, 3143-3149.

Lim, H., Carlton, A. and Turpin, B., *Isoprene Forms Secondary Organic Aerosol through Cloud Processing: Model Simulations*, Environ. Sci. Technol., 2005, 39, 4441-4446.

Goodwin, Bradley P., *Measurement of Ambient Black Carbon and Sulfate Aerosols at High Time Resolution*, Thesis presented to the Henry Edwin Sever Graduate School of Washington University in partial fulfillment of the Requirements of the degree of Master of Science, May 2005, Saint Louis, Missouri.

Schwab, J., Hogrefe, O., Demerjian, K., Dutkiewicz, V., Liaquat, H., Rattigan, O. and Felton, H., *Field and Laboratory Evaluation of the Thermo Electron 5020 Sulfate Particulate Analyzer*, Aerosol Science and Technology (2006), in press.

Appendix A: RAIN Instrumentation

Appendix A: RAIN Instrumentation

The RAIN sites employ the Sunset Laboratory OC/EC Carbon Aerosol Analyzers and Thermo Electron sulfate analyzers, which constitute the first use of these methods in routine, on-going state-run networks. These methods, combined with other historically routine measurements such as IMPROVE aerosol, NGN-2 (wet) nephelometers, continuous fine aerosol mass (PM_{2.5}), trace SO₂, ozone, meteorology, and automated digital visibility cameras (CAMNET), make up the RAIN core monitoring instrumentation.

The Model 3 OC/EC field analyzer from Sunset Laboratory uses an inlet cyclone to remove particles larger than 2.5 microns at an inlet flow rate of 8 LPM. The sample then flows through a carbon coated denuder to remove organic vapors. Particles are collected on a quartz fiber filter for approximately 100 minutes, and then selectively vaporized using a standard NIOSH 5040 protocol. Laser diode transmission is used to correct for pyrolysis of organic carbon. The resulting sample is oxidized to CO₂ using a MnO₂ oxidizer oven, and is analyzed via enhanced NDIR methods.

The Thermo Electron Corp. Model 5020 SPA sulfate monitor uses a BGI/Magee Scientific 0.732 sharp-cut cyclone inlet to remove particles larger than 1.8 microns at an inlet flow rate of 1.2 LPM. The sample then flows through a sodium carbonate coated denuder to remove SO_2 and then through an oven at 1000°C with a stainless steel rod, where sulfate is converted into SO_2 . The resulting SO_2 is analyzed via standard pulsed fluorescence methods. The system adjusts data for analyzer baseline drift and interference with an auto-zero cycle every 15 minutes.

All RAIN sites host IMPROVE samplers which operate on the every-third-day schedule endorsed by the U.S. EPA. The IMPROVE samplers at RAIN sites employ all four modules and supply highly detailed composition data on fine particles and some data on coarse ($PM_{2.5}$ to PM_{10}) particles. Ozone monitors at RAIN sites operate on ultraviolet photometry and meet EPA's "Federal Equivalent Method" specifications.

All RAIN sites track fine aerosol mass on a continuous, near-real time, basis. The Mohawk (CT) and Piney Run (MD) sites employ Met One model 1020 Beta Attenuation Monitors (BAMs) which detect accumulating aerosol mass via the drop off in intensity of a beta-ray beam passing through the sample. The McFarland Hill (ME) site employs a Tapered Element Oscillating Microbalance (TEOM) operated at 50°C for continuous fine mass measurement. The TEOM electronically monitors the natural harmonic frequency associated with a given mass to measure fine particle mass.

Appendix B: Data Validation and Caveats

Appendix B: Data Validation and Caveats

Carbon and SO₄ data collection began at RAIN sites during June or July (depending on the site) of 2004. Although the seasonal nature of fine aerosol chemistry in the Northeast cannot be over emphasized, this report covers data from site inception through the end of March 2005. Since the largest aerosol composition changes occur between summer and winter, we expect the three seasons (summer, autumn and winter) covered by this report will provide a reasonably complete view of annual data. Note, however, that at the time we completed this report, IMPROVE data were available only through December 31, 2004, which somewhat limited the inter-comparison of data.

Much of the instrumentation employed by RAIN is generic and has been field tested for many years. Some of the monitors, however, are state-of-the-art due to the burgeoning interest in fine aerosol composition. For example, the instruments that obtain sulfate (SO₄), organic carbon (OC) and elemental carbon (EC) data do so by converting aerosol mass to the gas phase (SO₂ and CO₂ respectively) before measurement takes place. The resulting gas-phase measurements are expressed in terms of standard temperature and pressure (25°C. and 29.92 in. Hg). Therefore, before analyzing or comparing to IMPROVE data, we had to convert the data to "local" conditions – a process that required on-site temperature data and an estimate of the site's average barometric pressure.

In addition to the STP-to-local processing, sulfate measurements might require a baseline "offset" which must be determined by periodic, on-site instrument checks. Similarly, OC measurements from the Sunset Laboratory monitor typically require subtraction of a "blank" value which, at this stage in instrument development, can only be generated by installing an appropriate filter on the instrument's inlet and operating it for at least 12 hours. At this time, a "universal" blank equal to 0.5ugC/m³ has been adopted for use in data assessment. This blank value may change and/or become site specific as knowledge of the performance of the Sunset instrument grows.

The reconstruction of both fine particle mass and of visual range requires additional processing to account for the lack of specificity in measurement and for the impact of water vapor on the size of particles. Relative humidity function factors range from ~2.5 to ~4.0 in the Northeast. The IMPROVE program has adopted a factor of 1.4 to convert measured carbon mass to an equivalent ambient organic carbon mass. In light of recent research (Hand and Malm, 2005), a factor of 1.8 has been used in this analysis for RAIN data. Unless otherwise indicated, all IMPROVE data expressed as organic matter in this memo will use the 1.4 factor.

Despite best planning efforts and assiduous operator attention, invalid data occur in all networks and RAIN is no exception. The collection process flags most invalid data but some suspect values are only noted during subsequent data analysis. In a smaller subset of cases, data may appear to be invalid in spite of supporting documentation showing its (apparent) validity. Each reporting organization has screened the RAIN data in this report at "Level 1" as per the RAIN Quality Assurance Project Plan and we have eliminated all invalidated data from that screening process. We will identify suspect values subsequently flagged during data analysis if they have been censored and, where feasible, identify the cause (or suspected cause) for their occurrence.

Appendix C: Sulfate Monitor Performance Issues

Appendix C: Sulfate Monitor Performance Issues

The Thermo Electron Model 5020 data used in this memo are consistently low compared to collocated IMPROVE filter sulfate 24-hour samples. The bias ranges between 56% and 74%, although correlation is very good with R^2 values ranging from 0.93 to 0.97. The highest regression coefficient ($R^2 = 0.97$) comes from the Piney Run site (Frostburg MD) which had the most bias in the Model 5020 sulfate data.

These low results are similar to what researchers at SUNY-Albany Atmospheric Sciences Research Center have observed (Schwab et al., *Aerosol Science and Technology*, 2006, in press), both in the field and the lab. One possible cause is poor conversion efficiency. However, two observations suggest that is not why the instrument is reading low, and it may be the instrument, not the method, that is the issue. In this discussion the terms "instrument" and "method" have different meanings; the instrument is the method as manufactured by Thermo Electron as the model 5020. The method is the fundamental technical basis upon which the instrument is built.

It has been observed at all three RAIN sites that there are cases that clearly show the "zero-cycle" signal ("F0 and F1" in the raw instrument cycle data output files) tracking the reported cycle sulfate concentration (the baseline analyzer response is subtracted from the gross measurement cycle response to provide the "net" response (sulfate). This should not occur unless there are pollutants such as NOx or H₂S (or other reduced sulfate species) that can cause a non-aerosol response in the instrument that covary temporally (on a 15-minute time scale) with the sulfate. SO₂ should not cause a baseline response, since it is removed with high efficiency upstream of the system. While this could occasionally happen, it should be rare. The clearest case of this kind of effect is from Piney Run during July 2005 (Figure C-1). This plot shows the reported 15minute sulfate on the x-axis, and the "baseline" (F0) signal (reported in ppb SO₂ by the analyzer) times 3.9 to make it equivalent to sulfate concentration (1 ppb SO₂ = 3.93 μ g/m³ sulfate at STP). The R² for these 15 minute data is 0.94, and the slope is 0.22 (22% of the signal that should be classified as sulfate is being lost by the baseline tracking the sulfate signal).

This very high correlation (given the short time intervals used here) over a wide dynamic range (1 to $28 \ \mu g/m^3 \text{ SO}_4$) makes a very convincing argument that there is something wrong with the instrument. To a lesser degree (in consistency but not magnitude), we see this across all three sites, and the 22% is consistent with the Schwab work. All this implies a basic design problem with the instrument, and not an isolated case of operator error.

There is strong evidence that the under-reporting of sulfate is not a fundamental problem with "the method" – e.g., poor converter efficiency or losses in the sample train, etc. An intensive evaluation and comparison of the method was carried out over a one-year period at the St. Louis EPA supersite by Jay Turner, Brad Goodwin, and others. The Model 5020 continuous sulfate method was compared to several different independent measurements of sulfate, both real-time (Particle Into Liquid Sampler - PILS) and 3 different filter-based integrated methods. In all comparisons, there was very good

correlation and numerical agreement between the Thermo method and the other methods. This work has been written up in detail in a master's thesis (Goodwin, 2005).

Figure C-2, Figure C-3 and Figure C-4 show the pre-production Thermo sulfate method (the same product used in RAIN sites) comparing well with PILS hourly sulfate data, 6-h, and 24-h integrated filter samples (IC analysis). R^2 values range from 0.95 to 0.97 (for the 1-h PILS data); regression slopes range from 1.00 to 1.04 (Thermo sulfate being the dependent variable), indicating that the Thermo method read equal to or slightly higher than the other methods. Similar results (slope = 0.98, R^2 =0.96) were observed when the Thermo method at the East St. Louis Supersite was compared to daily Speciation Trends Network sulfate measurements made at the St. Louis Blair St. site 10 km NW of the Supersite location (see Table C-1 for summary results of the four comparisons detailed by Goodwin).

The only significant difference in how the Thermo method was operated in St. Louis for these comparisons was that the 10-second raw data output from the analyzer was used and post-processed into 15 and 60 minute values -- in effect bypassing all the computational algorithms built into the instrument's firmware and used for computing all the rain sulfate data. We are investigating the possible causes of this difference with the manufacturer with the goal of correcting the instrument's production or configuration to minimize this bias. Limited evidence from other networks indicates that there may be problems with the SO2 instrument's response time. This would explain the observed bias and tracking of the baseline signal with the sulfate data. At one non-RAIN site, the majority of the low bias was resolved when the SO2 detector of a 5020 sulfate analyzer was replaced.

One additional question is why the Piney Run (Frostburg MD) Thermo sulfate results were notably lower than the other two RAIN site comparisons. One possibility is that there are some losses when the aerosol is strongly acidic (approaching sulfuric acid). That would typically be the case only when sulfate levels were high. There is limited evidence at that site that the Thermo sulfate response compared to filter sulfate is somewhat suppressed at higher levels, but no clear conclusions can be made from the data to-date. Robust lab tests with sulfuric acid aerosol on the present instrument configuration have not been done, so this remains an open question.

Finally, it should be noted that even though there is a substantial bias between the filter and continuous data, the high degree of correlation at all RAIN sites allows us to correct (linearly scale) the hourly sulfate data to be equivalent to the IMPROVE filter sulfate data.

Figure C-1: Sulfate Concentration vs. Baseline Signal (times 3.9)





Figure C-2: Response (1-hour) of Model 5020 vs. PILS (1-hour) results. (Goodwin 2005)



Figure C-3: Model 5020 1-hour SO₄ results vs. 6-hour Mini-HEADS results. (Goodwin, 2005)



Table C-1:	Regression statistics for pre-production sulfate unit compared to other
	sulfate quantification methods. (Goodwin, 2005)

Comparison	Duration	Slope	Intercept	\mathbf{R}^2	Ν
HEADS	24 Hour	1.04 ± 0.10	-0.15 ± 0.41	0.95	25
Blair St.	24 Hour	0.98 ± 0.04	-0.15 ± 0.21	0.96	96
MiniHEADS	6 Hour	1.00 ± 0.04	-0.01 ± 0.23	0.95	106
PILS	1 Hour	1.04 ± 0.01	0.04 ± 0.08	0.97	514