Evaluation of Filter Recovery Period for the Determination of Fine Particulate Matter as $PM_{2.5}$ in the Atmosphere

M. Papp, S. Eberly, T. Hanley, N. Watkins,[1] H. Barden[2], G Noah,[3] R. Bermudez, R. Eden[4]
B. Franks,[5] A. Johnson, R. Marriner, [6] E. Michel[7]

ABSTRACT

In July of 1997 the Federal Reference Method (FRM) for $PM_{2.5}$ was promulgated to provide a measurement system for monitoring as part of the new form of the fine particulate matter (PM) standard. In order to ensure the quality of the resultant concentration data, a number of quality control requirements for field and laboratory activities were developed in the Code of Federal Regulations (Parts 50, 53 and 58) and in various guidance documents (e.g. Guidance Document 2.12). The requirements and guidance were developed to ensure a consistent approach to data collection activities and to provide data of adequate quality to support the decision making process. During the first year of $PM_{2.5}$ monitoring, the State, Local and Tribal monitoring agencies identified the requirement to collect samples within 96 hours of the end of the sample period as burdensome with an undetermined amount of value for ensuring the quality of the data. A workgroup of stakeholders agreed to conduct a study designed to determine whether filters recovered 7 days after the end of the sample period meet the precision and bias data quality objectives (DQOs) of the $PM_{2.5}$ FRM measurement system. The study was initially performed in Research Triangle Park, North Carolina and replicated in 5 other locations around the nation to provide spatial representation. Precision and bias data from all study sites met the DQOs leading to the conclusion that extending the filter retrieval period did not significantly affect data quality.

[1] U.S. EPA Office of Air Quality Planning and Standards

[2] U.S. EPA Region 4

[3] Integrated Laboratory System

[4] South Coast Air Quality Management District, CA

[5] Puget Sound Clean Air Agency, WA

[6] Maine Dept. of Environmental Protection

[7] Texas Natural Resource Conservation Commission

INTRODUCTION

The reference method for the determination of fine particulate matter as $PM_{2.5}$ in the atmosphere is defined by regulation in Appendix L of 40 CFR Part 50. Within the reference method there are many requirements for laboratory and field data collection activities, one of which is the time allowed for the recovery of a sample once sampling is completed. Section 10.10 of Appendix L states: "Within 96 hours of the end of the sample period, the filter, while still contained in the filter cassette, shall be carefully removed from the sampler,..." The mass on a filter left in a $PM_{2.5}$ FRM sampler after the end of the sample period can change due to a number of factors that may result in increase and/or losses in filter mass. Contamination can occur as filters are not perfectly protected from fugitive dust or aerosols. The mass on a filter can increase due to adsorption or decrease due to volatilization of gases and/or vapors from the particles on the filter.

During meetings to discuss the implementation of the $PM_{2.5}$ monitoring, the filter recovery requirement of 96 hours was identified as one of the top resource burdens to implementing the network and a burden for which an undetermined amount of value might be gained. From a monitoring agency's perspective, the 96-hour retrieval requirement means each sampler must be visited nearly two times per week If the sample collection period could be increased from 96 hours (4 days) to 168 hours (7 days), each sampler could be visited once per week, which translates into a significant decrease in burden on the monitoring agencies.

In order to determine whether extending the filter recovery period affects the quality of the data, a study was designed to answer the question: Does a sample collected up to 7 days after the end of a sample period meet the data quality objectives (DQOs) of the $PM_{2.5}$ FRM measurement system? The DQOs for the $PM_{2.5}$ Network are 10% precision (measured as a coefficient of variation) and $\pm$ 10% bias. Precision is estimated using collocated monitors of like method designation. Since $PM_{2.5}$ is a manual method and there are presently no known performance evaluation samples or standards, bias is estimated using the Performance Evaluation Program (PEP). The PEP program is implemented nationally using certified field scientists who set up an FRM monitor at a routine monitoring site, collect a sample and send this sample to one of two independent PEP laboratories. Bias is determined by comparing the routine and PEP values. More information on the PEP (Implementation Plan, SOPS and QA Project Plan) can be found on the AMTIC $PM_{2.5}$ website (http://www.epa.gov/ttn/amtic/pmqainf.html).

This report describes the study design and the results that will help to determine whether extending the sample retrieval period will meet the current $PM_{2.5}$ DQOs and therefore produce data of acceptable quality.

STATEMENT OF PROBLEM

The $PM_{2.5}$ mass DQOs have the goal that the bias be between -10% and +10% and that precision be less than 10%. Bias is estimated using percent difference; precision is calculated using the coefficient of variation (CV). If we can show that retrieving a sample after 7 days does not result in a violation of the

bias and CV DQOs, then it would be reasonable to allow the retrieval time to be increased to 7 days. Evaluating any violation of the DQOs will require the estimation of the bias and precision Estimation of the CV will require multiple samplers of similar method designation (as defined in CFR) and estimation of bias will require comparison to PEP samplers (as defined in CFR). Although the study seeks to determine whether a sample that was 7 days old in a sampler does not result in a violation of the bias and CV DQOs, it should be noted that this is a conservative test of the question. The reason for this is that samples that are 7 days old represent the maximum exposure time in a sequential sampler (most common type in current $PM_{2.5}$ network). Whether on the one-in-three day or daily sample schedule, some samples would be up to 7 days old while others would be a few days old. Therefore, while the study examines whether a sample up to 7 days old still meets the DQO's for CV and bias, if relief is granted and implemented, most filters would still fall into the existing criteria for sample recovery (4 out of 7 filters for daily sampling, and 2 out of 3 filters for the one in three day sampling).

STUDY DESIGN

This section describes the rationale for the study duration, location of the sampling platforms, the sample design for each platform and the data evaluation techniques.

Study Duration:

The study was designed to collect data for an entire year in order to make a decision that the DQOs are met, although it may be possible to prove that the DQOs are violated with less data. Reasons for collecting a year's worth of data included capturing:

- seasonal differences due to meteorological factors of temperature and humidity
- differences on the chemical composition on the filter (large percentage of nitrate and organics vs. non-volatiles)
- the spectrum of operating conditions

Number of Samples:

Each platform that participated in the study should have at least 20 sampling events for the year with a goal of 5 in each quarter. More information would improve the confidence in the results.

Sample Locations:

It was a study objective to locate sampling platforms in parts of the country that were likely to experience different meteorology and be exposed to areas with varying chemical composition. For the intercomparison study for the $PM_{2.5}$ speciation samplers, the locations identified to represent the spectrum of chemical composition were Rubidoux, CA (very high nitrate, moderate organic material, and low sulfate and crustal material), Philadelphia, PA (high sulfate and organic material but low nitrate), Phoenix, AZ (high crustal material and strong nitrate and organic material component), and Research Triangle Park, NC (long-term site at lower concentration range). The second phase of the

intercomparison study was to include Utah or Washington state to sample under conditions with high wood smoke emissions, and/or Atlanta to sample under conditions with high biogenic carbon emissions. Thus the locations for this study should minimally be in CA, the northeast, the arid southwest, Washington/Utah/Idaho, and the Atlanta area.

OAQPS asked for volunteers from the State, Local and Tribal organizations to participate in this study. Table 1 provides a listing of the organizations participating in the study and the sampling equipment used. It is felt that the study sites provided adequate coverage of varying meteorological and compositional differences.

Design for Each Sampling Platform:

This section describes what equipment should be on each platform (site) that participated in this study, how the equipment should be operated, and how the filters should be handled. Figure 1 provides a time line of the sampling events.

(1) At each monitoring platform, place two 48-hour samplers (48-hour sample retrieval period) and three 177-hour samplers (177-hour retrieval) that are of the same method designation to ensure the collection of at least one 48-hour sample and two 177-hour sample values for a particular day. All of the samplers must be operated on the same days for the study.
(2) Operate each 48-hour sampler and handle each filter used in the 48-hour samplers exactly like they are being used to perform an evaluation of one of the monitors in the primary $PM_{2.5}$ network, that is, according to the PEP QA Project Plan (QAPP). This means that the samplers are verified for each sample run (one-point checks) and that the Labs in Region 4 or 10 are pre- and post-weighing the filters. The 48-hour filters should be retrieved within the time allowed in the PEP QAPP, meaning within 48 hours after sample collection.
(3) Operate each 177-hour sampler and handle each filter used in the 177-hour samplers exactly like they are part of the primary $PM_{2.5}$ network, that is, according to the state's QAPP. This means that the monitors are operated by the people identified in the state's QAPP and the filters are pre- and post-weighed by the lab identified in the state's QAPP. The 177-hour filters should be retrieved at about 177-hours (~9 am on the 8^{th} day).
(4) For a day's data to be used in this experiment, at least one of the 48-hour monitors must have a sample that meets all of the PEP validation criteria and at least two of the 177-hour monitors must have samples that meet all of the 177-hour validation criteria. Days that meet this requirement will be called **sampling events**.

Table 1. Filter Extension Study Location, Contact(s) and Monitoring Equipment

State/ Site	Contact(s) and Organization	48 hr. sample retrieval	177 hr. sample retrieval
CA Rubidoux	Rene Bermudez & Rudy Eden South Coast AQMD	Andersen Portable (2)	Andersen Sequential (2)
GA Athens	Herb Barden & Greg Noah EPA Region 4	BGI Portables (2)	R & P Single (3) Andersen Single (2)
ME Augusta	Andy Johnson & Rick Marriner Maine Dept. of Environmental Protection	Andersen Portable (2)	R & P Single (1) R & P Portable (1)
NC RTP	Tim Hanley & Nealson Watkins EPA OAQPS	Andersen Single Channel (2) BGI Single Channel (2) R&P Single Channel (2)	BGI Single Channel (1) Andersen Sequential (3) R & P Sequential (2)
TX Austin	Ed Michel Texas Natural Resource Commission	Andersen Portables (2)	R & P Sequential (3)
WA Seattle	Bob Franks Puget Sound Clean Air Agency	Andersen Portables (2)	R & P Sequential (3)

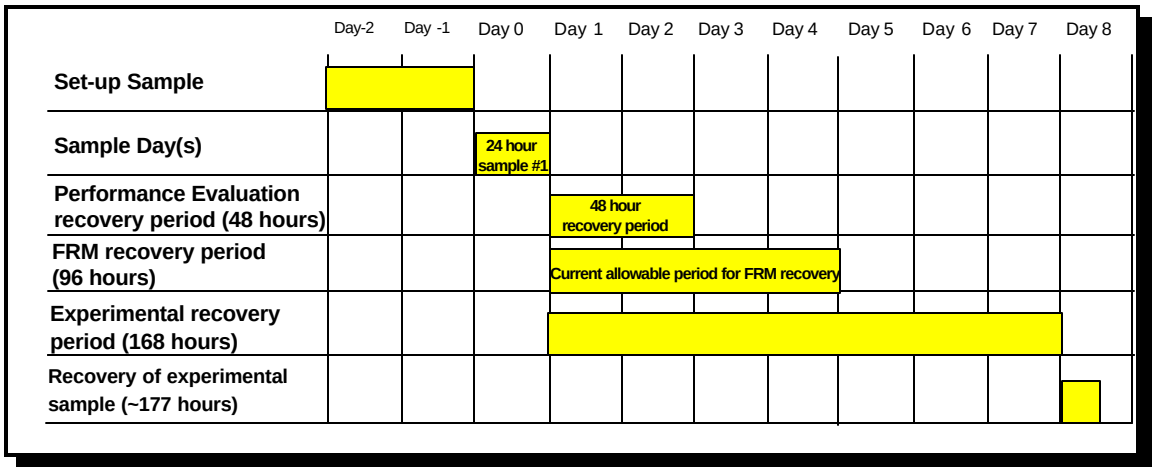


Figure 1 Study retrieval periods.

Data Evaluation

Each method designation will be analyzed separately. The analysis will be comprised of a couple of steps. First it will be determined whether the 177-hour samplers and the 48-hour samplers are generating data that are sufficiently consistent. If both the 177-hour samplers and the 48-hour samplers are sufficiently consistent, then a measure of bias between the instruments will be estimated. For the ambient air monitoring program, precision and bias comparisons are not performed if any concentration value for a sampling event is below 6 $\mu g/m^3$. For this study some of the evaluations will include values less than 6 $\mu g/m^3$, however, final results and conclusions will be made from the data set containing sampling events in which all concentrations were greater than 6 $\mu g/m^3$.

Three step analysis:

(1) **Determine whether the replicate 177-hour samplers are producing results that are consistent.** To do this, we estimated the CV for each day and aggregated the estimates for all observations collected at a given site. The formula (equation 1) to estimate the squared coefficient of variation for site *i* is:

$$CV_i^2 = \frac{1}{n}\sum_{t=1}^{n}\left[\frac{\sum_{r=1}^{s}\left(X_{itr} - \overline{X}_{it}\right)^2 \Big/ (s-1)}{\left(\sum_{r=1}^{s} X_{itr} \Big/ s\right)^2}\right] \qquad \textbf{Equation 1}$$

Heuristically, this estimator is the average squared coefficient of variation, where the average is taken over the *n* sampling periods. The coefficient of variation for each sampling period (what is inside the square brackets) is estimated by the sample standard deviation of the *s* samplers divided by the sample mean of the *s* samplers.

(2) Determine whether the replicate 48-hour samplers are producing results that are consistent. To do this, we estimated the CV for each day and aggregated these estimations for all observations collected at a given site. Since there are only 2 48-hour samplers, the formula above can be simplified and the simplified form matches that in 40 CFR Part 58 Appendix A, Equation 21. Specifically, the simplified formula (equation 2) is:

$$CV_i^2 = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{\left(X_{it1} - X_{it2}\right)\Big/\sqrt{2}}{\left(X_{it1} + X_{it2}\right)\Big/2}\right)^2 \qquad \textbf{Equation 2}.$$

Note that significantly large estimates of coefficient of variation can be due to (1) the instruments producing concentrations that are highly variable, (2) one instrument being consistently higher or lower than the other(s), or (3) a combination of the two. That is, the CV estimate includes precision and relative bias between the two instruments..

(3) Determine whether there is a unacceptable percent difference between the concentrations produced by the 177-hour samplers and the 48-hour samplers. We compared the average concentration from the 177-hour samplers to the average concentration from the 48-hour samplers. This will be performed using equation 26 and 27 from 40 CFR Part 58 Appendix A. Equation 26 calculates the accuracy of a single check. Of special note is that the 48-hour samplers are represented by X_t and therefore only the 48-hour samplers are used in the denominator since they are the baseline of the comparison. Equation 26 is:

$$d_t = \frac{Y_t - X_t}{X_t} \times 100\% \qquad \textbf{Equation 26}$$

Where Y_t is the 177-hour average concentration and X_t is the 48-hour concentration value for day t.

Once each of the individual checks are known the average of the individual percentages during the study for site i can be calculated by equation 27. Equation 27 is:

$$D_i = \frac{1}{n} \sum_{t=1}^{n} d_t \qquad \textbf{Equation 27}$$

STUDY IMPLEMENTATION

The study was implemented in two phases. In August of 1999, OAQPS implemented the study at the Research Triangle Park Air Training Facility (ATF) platform. Since the ATF had a number FRM samplers, it was cost efficient to start the study at one site and if the data proved promising, to implement the study at the remaining monitoring sites. By April of 2000 the study participants felt the ATF data looked promising and started the process of acquiring the necessary monitors at the sites identified in Table 1. Monitors were loaned and borrowed from various monitoring organizations and the PEP program to make up the compliment of monitors shown in
Table 1.

The design for each sampling platform called for:

- **Two 48-hour monitors and three 177 hour monitors** to ensure that we would have at least one 48-hour sampler and two 177-hour sampler to make up a sampling event. This was accomplished for four out of the six study sites. The study was not able to provide a third 177-hour monitor to the CA and ME participants.
- **Operate the 48-Hour samplers similar to PEP QAPP and the 177-Hour samplers similar to the State QAPP.** This was accomplished with the 48-hour filters being pre-weighed and post weighed by one of the two national PEP laboratories and implementing the PEP samplers as called for in the PEP QAPP. The only deviation allowed in the PEP QAPP was the frequency of the one point verifications (flow rate, temperature, barometric pressure). The PEP QAPP calls for these verifications to be performed at every sampling, due to the fact that the portable PEP samplers are normally set up, taken down and transported after every sampling event. In the case of the Filter Extension Study, the portable samplers were permanently set up at the study sites and therefore the verifications were performed based upon the States QAPP requirements.
- **A valid sampling event of one valid 48-hour value and two valid 177-hour values**. In general, this was met, but is discussed in more detail in the results section.

Sampling at the participant monitoring sites got underway at somewhat different time periods, based on arrival of equipment, and time to become familiar on FRM equipment that were different from the FRM equipment the volunteer organizations operated. Table 2 provides the sampling start and end dates and the various concentration ranges collected during the study period.

Table 2. Study Initiation/Completion

State/Site	Start Date	End Date	Concentration Ranges(ug/m³) 48-Hour Values		
			~5% Conc.	Mean	~95% Conc
CA/ Rubidoux	11/05/00	12/09/01	9.47	32.26	65.69
GA /Athens	07/17/00	10/23/01	7.60	18.84	30.16
ME/Augusta	09/03/00	09/19/01	6.38	14.02	20.94
NC/RTP	08/28/99	08/23/00	6.75	14.51	26.85
TX/Austin	10/31/00	10/02/01	6.75	13.23	20.25
WA/Seattle	08/08/00	08/21/01	7.36	17.27	40.9

RESULTS

This section discusses the study data in relation to the achievement of the completeness, precision and bias data quality objectives. As mentioned earlier, precision and bias estimates are usually not made on sample events where concentrations below 6 $\mu g/m^3$ are reported. Some of the results for this study will be reported with and without data from sample day/sites with concentration values less than 6 $\mu g/m^3$. However, final conclusions will be drawn from estimates with values less than 6 $\mu g/m^3$ removed.

A 177-hour sample concentration value of 6.17 $\mu g/m^3$ from the CA study site was removed from the database. 48-hour values for this sampling date were 24.09 and 23.42 $\mu g/m^3$ and the remaining 177-hour value was 22.96 $\mu g/m^3$. This one outlier changed the 177-hour precision estimate from 6.88 to 17.88 and would also have had an effect on the bias estimates for the sampling date and the study site average. In addition, for the entire data set, there were no 48 and 177-hour concentrations with a difference of greater than 5.25 $\mu g/m^3$. Although there was no information available to invalidate this sample, the corroborating information from the other 3 concentration values led the authors to conclude the sample concentration was an anomaly and removed it from the evaluation.

Sampling Event Completeness

A sampling event was defined as a day where one valid 48-hour sample and two valid 177-hour samples are collected. The study design had the objective of collecting 20 valid sampling events for each study site with a distribution of 5 per quarter. Table 3 provides the completeness results for each study site based on the definition of a sampling event. The values in parentheses in some of the quarterly results represent the number of days where there was only 1 valid 177-hour sample retrieval data point to compare with the 48-hour sample retrieval data point. The last column of Table 3

represents the number of days from each study site where there are at least one valid data point for both 48 and 177-hour sample retrieval times. Although the information is not reported, there were a number of days where either the 48-hour or 177-hour samples where determined to be invalid which reduced the completeness for that study site for the bias estimate. However, precision data for these incomplete sampling events could be used. Since the sampling end dates for the study occurred in either the 3rd or 4th quarters (see Table 2) there was no possibility for additional sampling in incomplete quarters without extending the study into 2002.

Due to the voluntary nature of the study, the time frame for decision making, and a review of the results, the study participants did not feel it necessary to delay data evaluation for the collection of additional samples. Since there were two sites that did not have the advantage of a third 177-hour monitor, the data represented in the last column will be used in subsequent precision or bias evaluations. With the exception of Texas, the study sites did collect 20 or more samples and 63% of the quarters had 5 or more valid samples.

Table 3. Filter Extension Study Site Completeness Statistics.

Study Site	Quarter 1 (Jan-Mar)	Quarter 2 (Apr-June)	Quarter 3 (Jul-Sept)	Quarter 4 (Oct-Dec)	Total (Initial SD[1])	Total Valid 48/177hour
CA/ Rubidoux	0 (2)	6(4)	7(1)	12 (2)	25	34
GA /Athens	4	6	11	8	29	29
ME/Augusta	2 (3)	5 (1)	6	4 (3)	17	24
NC/RTP	3 (1)	7	14 (1)	4	28	30
TX/Austin	4 (1)	3	1	5	13	14
WA/Seattle	5	6 (2)	11	8 (3)	30	35

1 =Based on study design requiring 2, 177-hour sample pairs per sampling event.

Precision

Figure 2 provides a summary of the precision estimates of the 48-hour and 177-hour retrieval times. The variables listed as "48 Hour-6" and "177 Hour-6" refer to sampling events where all the sample concentrations were above 6 $\mu g/m^3$. In addition, the State collocated precision estimates that are reported in the CY2000 $PM_{2.5}$ QA Report are also provided as a comparison. The NC and GA sites do not have these values reported because they were operated by EPA and were not part of a state monitoring network. Figure 2 also provides national precision estimates (variables CY99 and CY00) as reported in the CY1999 and CY2000 $PM_{2.5}$ QA Reports respectively. The CY1999 and 2000 results and the State-CY00 results are derived from samples that where retrieved anywhere from 8 to 96 hours and are provided for comparison purposes only. Precision estimates for the study sites are also summarized in Table 4.

The statistics used to estimate precision can be affected by the number of samples and low concentrations. One or two "relatively large" imprecision values can have a significant effect on

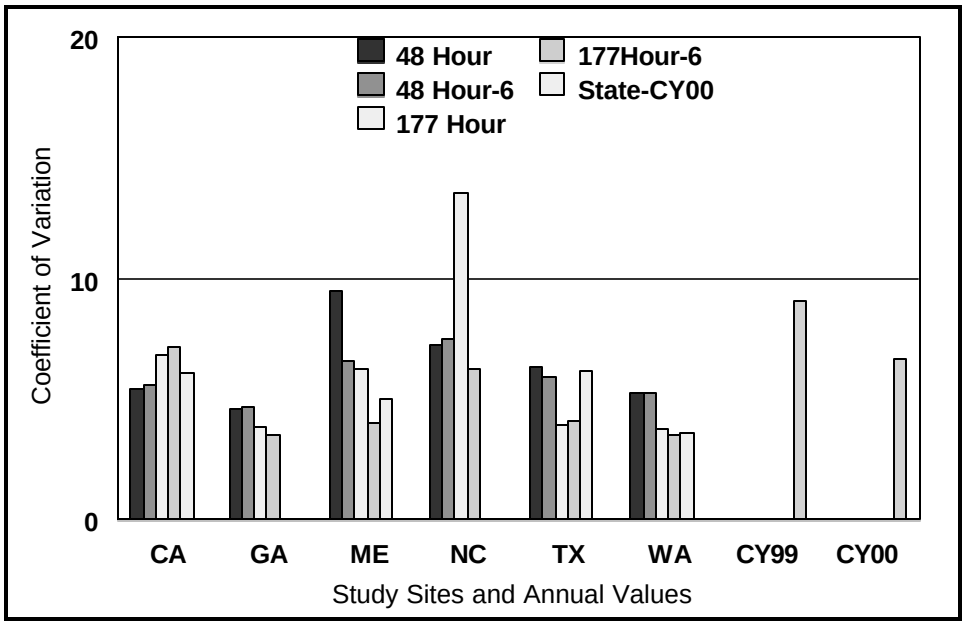


Figure 2. Filter Extension Study precision results

precision. In addition, sites or samples with low concentrations have a tendency for greater imprecision since a small absolute difference in a sample pair can produce large coefficients of variation, especially when the coefficient of variation is squared (using equation 1 above). Two pairs of values significantly affected the precision values for the NC and ME Sites:

< One pair (8.27 and 3.04 µg/m^3) changed the 177-hour NC site mean precision estimate from 6.04% (w/o pair) to 13.59 % (with pair). Since the other sample concentrations at this site on this sample date were below 6 µg/m^3, the precision estimate for the "177-hour-6" sample estimate did not include this high CV and therefore did not effect the mean.

< One low concentration pair (2.83 and 4.17 µg/m^3) changed the ME 48 hour precision estimate from 6.16% (w/o pair) to 9.18 % (with pair). Since both sample pairs were below 6 µg/m^3, the precision estimate for the "48 hour-6" sample estimate did not include this high CV.

In most cases, using samples with concentrations less than 6 µg/m^3 had very little effect on the overall precision estimate (with the exception of ME and NC). Figures 3 and 4 provide box and whisker plots of the 48-hour and 177-hour precision estimates stratified by study site. The plots provide median (middle line), mean(+), 25th and 75th percentiles (top and bottom of box), and the 5th and 95th percent of the data (end of whisker). The results in the two figures do not use any sampling events where values were < 6 µg/m^3.

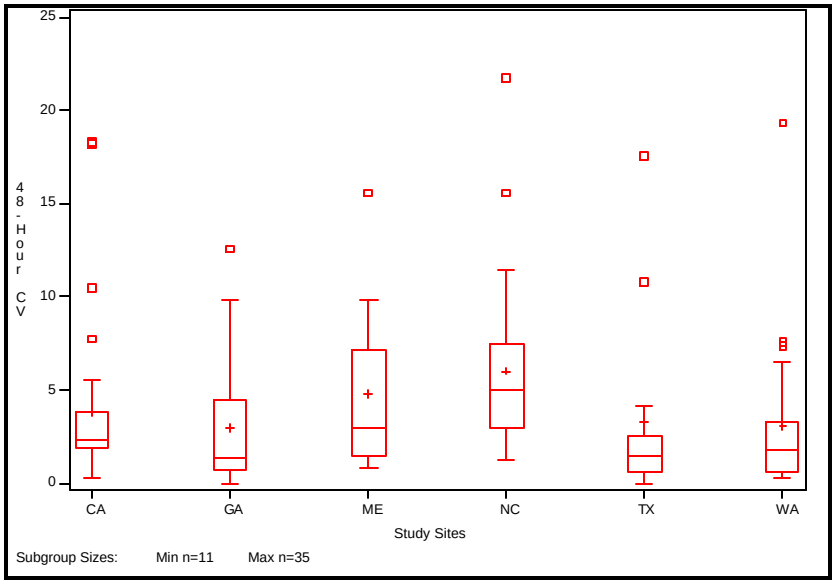


Figure 3. Box and whisker plots of the 48-hour precision estimates.

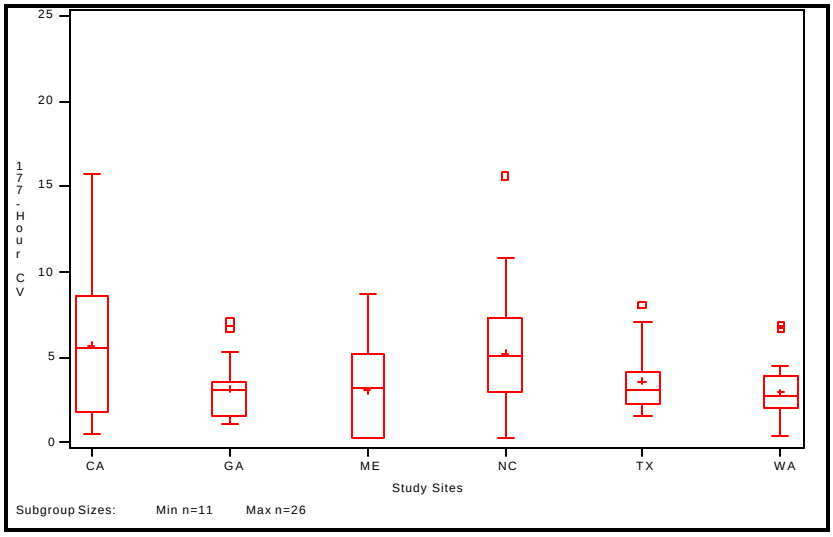


Figure 4. Box and whisker plots of the 177-hour precision estimates.

The results in Figures 3 and 4 provide slightly lower estimates of mean precision by study site than are reported in Table 4 because the results in the figures are not squared as described in equation 1 above. However, the figures can be used to illustrate the repeatability within and between study sites and filter recovery periods.

In general, the precision results are comparable to the State CY2000 precision results (see Table 4). Since the national precision estimates are generated by first removing all sample pairs with values less than 6 $\mu g/m^3$, comparing both years (99 and 2000) against the study data that removes the values less than 6 $\mu g/m^3$ illustrates that most of the study sites precision estimates were comparable or better than the national averages and the overall study mean precision estimates for both sample retrieval periods are below the overall national precision results reported in the CY1999 and CY2000 $PM_{2.5}$ QA Reports. Figure 5 provide a cumulative distribution summary of the precision CV results for all the study sites which indicate that over 90% of all the precision estimates met the 10% CV DQO goal.

It was concluded that both the 48-hour samplers and the 177-hour samplers achieved the precision DQO goals and were consistent within and between sites so that bias between the two sample retrieval times could be evaluated.

Table 4. Precision Results

Org	48 Hour CV / n	48 Hour CV - 6 /n	177-hour CV / n	177-hour CV - 6 / n	CY99/00 QA Report
CA	5.39 / 38	5.54/ 35	6.88 / 26	7.19 / 23	– / 6.1
GA	4.54 / 25	4.53 / 23	3.84 / 29	3.52 / 26	– / --
ME	9.18 / 17	6.52 / 11	6.24 / 18	4.02 / 12	– / 5.0
NC	7.21 / 30	7.43 / 28	13.59 / 28	6.18 / 26	– / --
TX	6.33 / 14	5.95 / 13	3.94 / 13	4.09 / 13	– / 6.5
WA	5.25 / 28	5.24 / 21	3.77 / 30	3.52 / 23	– / 3.6
Totals	6.20 / 152	5.92 / 131	7.53 / 144	5.08 / 123	9.1 / 6.7

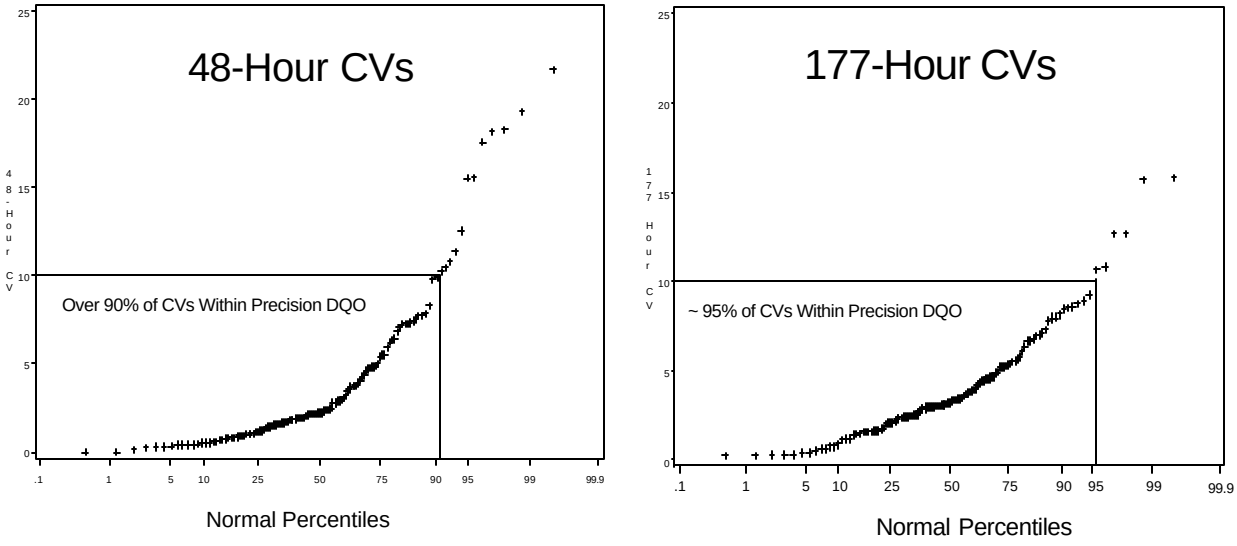


Figure 5. Cunulative distribution frequencies of the 48-hour and 177-hour coeiffients of variation (CV)

Bias Results

Bias is evaluated at each study site as well as aggregated. Table 5 and Figure 6 provide a summary of the bias estimates by study site. In addition, the State bias estimates that are reported in the CY2000 $PM_{2.5}$ QA Report are also provided as a comparison. The NC and GA sites do not have these values reported because they were operated by EPA and were not part of a state monitoring network.

Table 5. Filter Extension Study Bias Summary Statistics

Org	% Diff / pairs	% Diff - 6 / pairs	% Diff CY00 QA Report
NC	1.06 / 30	0.44 / 28	
GA	-1.11 / 29	-1.76 / 26	
CA	1.19/ 34	0.28 / 31	-0.8
WA	-5.18 / 35	-5.45 / 26	-3.0
ME	3.12 / 24	-2.02 / 16	1.5
TX	-7.99 / 14	-8.40 / 13	-12.2

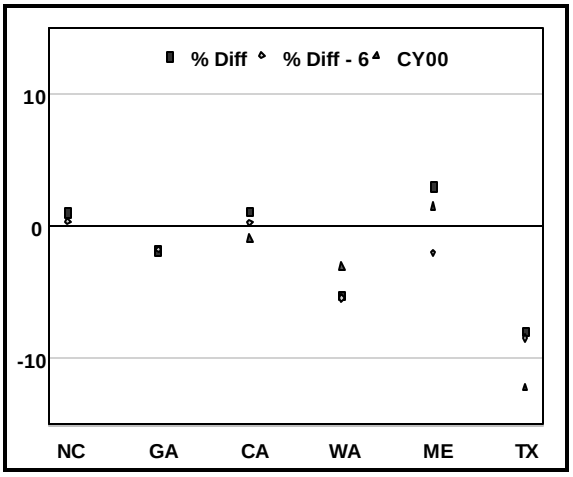


Figure 6. Filter Extension Study bias summary

Since the bias results with and without the use of values with concentrations less than 6 $\mu g/m^3$ are fairly similar, the remaining graphics and site summaries will be produced **without** sample days that had sample concentrations less than 6$\mu g/m^3$. Figures 7 through 12 provide bias summaries of the individual study sites.

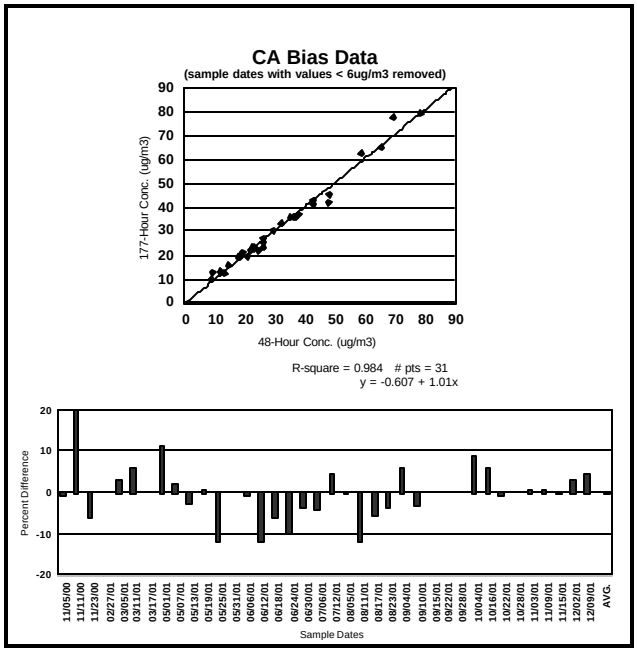


Figure 7. CA bias data

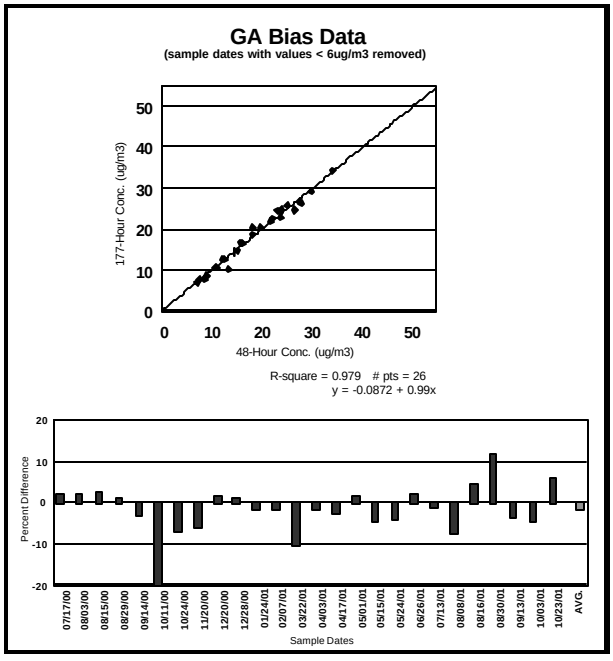


Figure 8. GA bias data

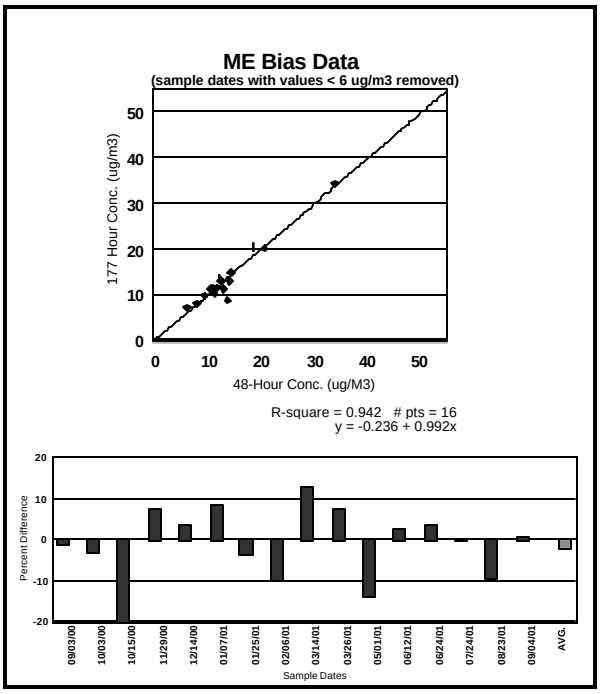


Figure 9. ME bias data

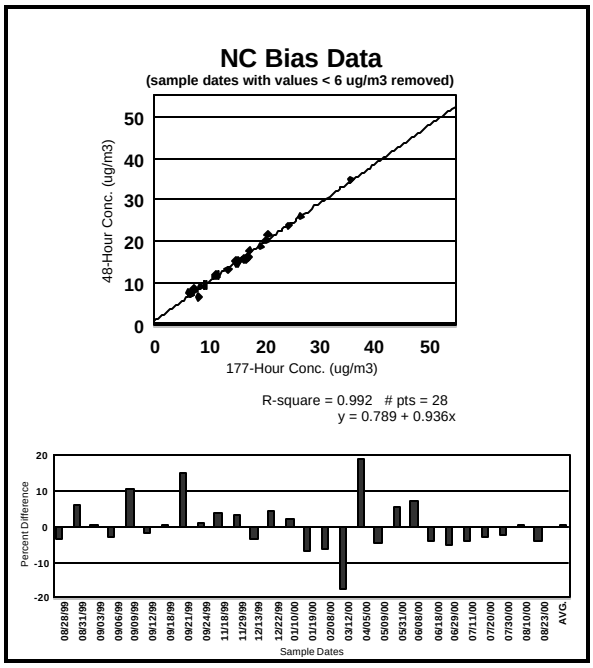


Figure 10 NC bias data

Four of the six study sites (Figures 7 through 10) produced bias results that were fairly random, meaning there did not appear to be a systematic difference between results reported at 48 hours and those reported at 177 hours. The WA and TX data had the largest bias estimates and the results appear to show some systematic bias across all sampling seasons. However, as is shown in the national estimates from the CY2000 QA Report (Table 5) the overall bias estimates are similar. Both the WA and TX sites were within the data quality objective requirements for the $PM_{2.5}$ program.

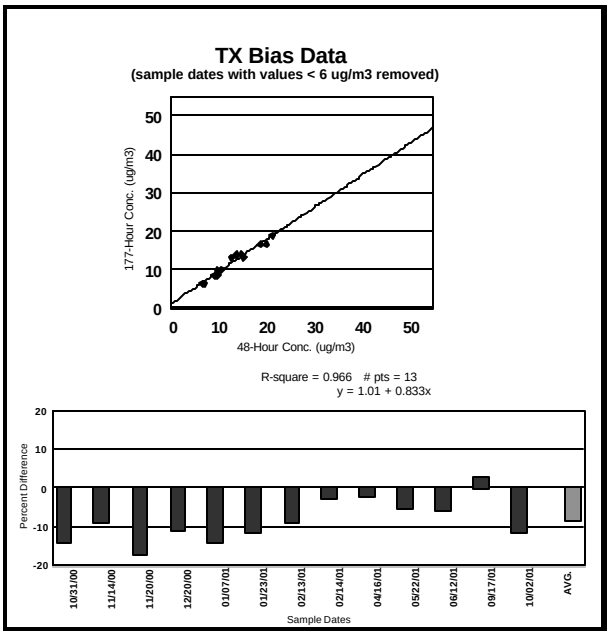


Figure 11. TX bias data

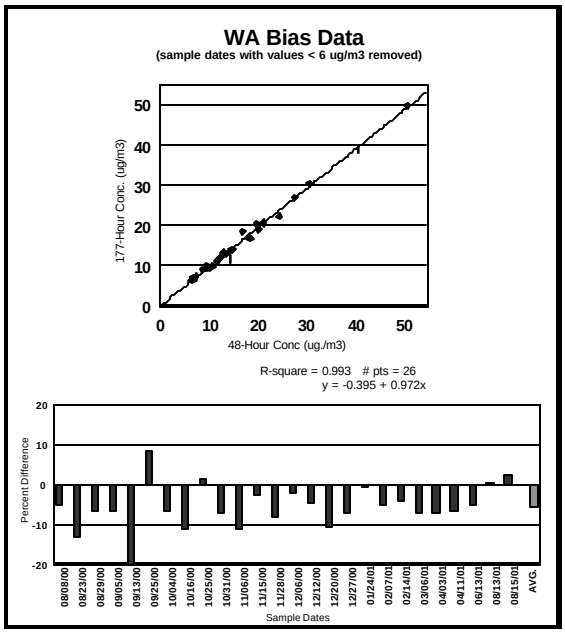


Figure 12. WA bias data

Since the NC and GA sites operated additional samplers of varying method designations, we also evaluated the bias at these sites by method designations to ensure that the method designations bias results were not being masked (e.g. a positive bias for one method designation balancing out a negative bias by a different method designation.) Similar to the results reported in Figures 7 and 10, the bias results for each method designation for GA and NC were similar in percent difference and direction.

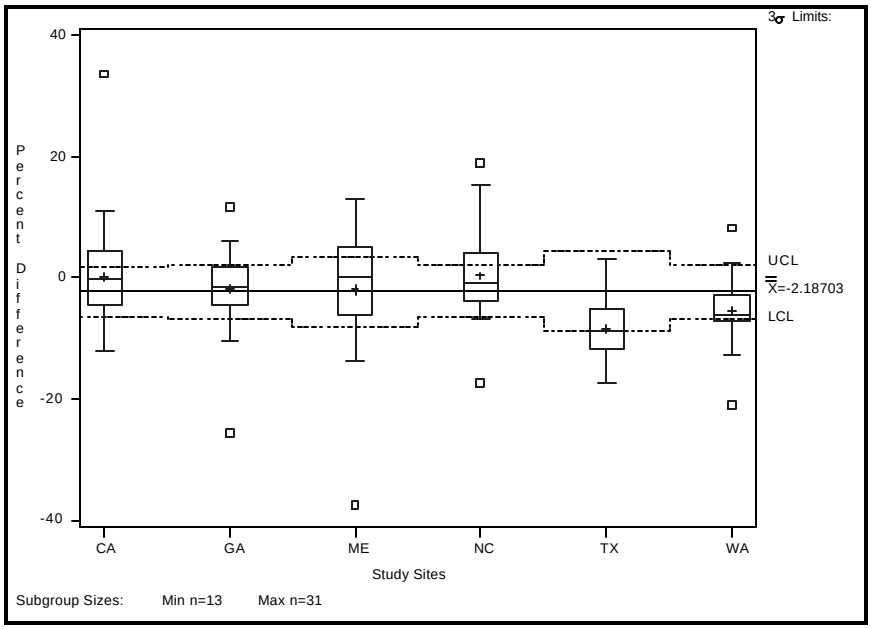


Figure 13. Box and whisker plots of bias estimates

The bias estimates have met the data quality objectives for all sites involved in the study. In addition, the study results seem to be comparable in direction (positive/negative) and percent difference to state bias values reported in the CY2000 QA Report. Similar to the box and whisker plots for precision, Figure 13 provides a summary of the bias estimates by study site. This plot also shows the overall mean percent difference as -2.2 % and demonstrates the influence each study site had on the upper and lower confidence intervals of the mean percent difference. This evaluation reveals that the bias estimates were very consistent within and between study sites.

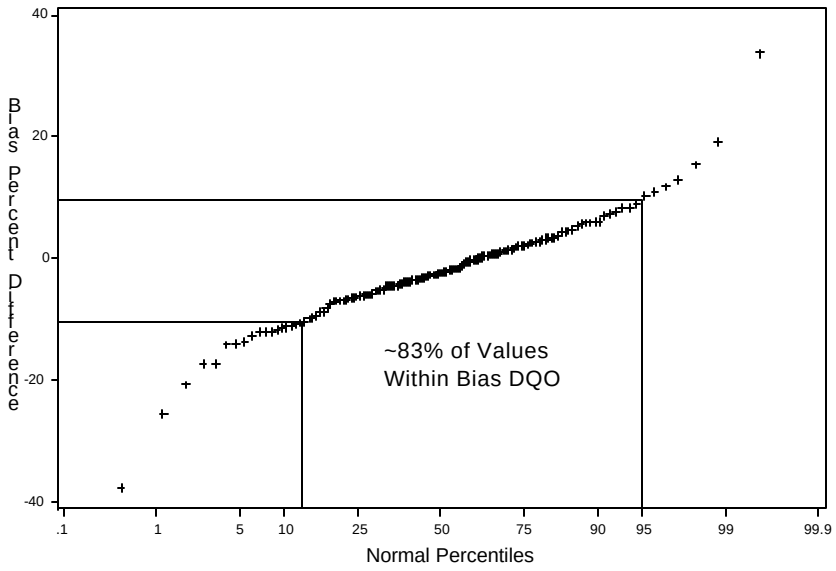


Figure 14 provides a frequency plot of the individual percent differences from all the sites. 83% of these individual comparisons were within the ± 10% DQO.

Figure 14 Cumulative frequency plots of bias estimates

Supporting Information

Table 6. Michigan Study Experimental Design

Days of Lag before Filter Pick up				
All Filters picked up between 7 and 7:30 a.m. EST				
Run Date	Monitor 1	Monitor 2	Monitor 3	Monitor 4
8/19/99	1	2	3	4
8/22/99	1	2	3	4
8/25/99	1	2	3	4
8/28/99	1	2	3	4
8/31/99	1	2	3	4
9/3/99	1	2	3	4
9/6/99	1	2	3	4
9/9/99	1	2	3	4
9/12/99	1	5	6	7
9/15/99	1	5	6	7
9/18/99	1	5	6	7
9/21/99	1	5	6	7
9/27/99	1	5	6	7
9/30/99	1	5	6	7
10/3/99	1	5	6	7

In August and September of 1999 the Michigan Department of Environmental Quality staff (courtesy of Mary Ann Heindorf) ran 4 collocated samplers in Lansing, MI to determine any influence of sample retrieval times on sample concentrations. In this study (see Table 6) filters were picked up 7 hrs, 31 hrs, 55 hrs and 79 hrs after completion of the sampling event or just shy of 1,2,3,4 days later. Halfway through the study, the study was repeated keeping the day 1 sample retrieval while the remaining 3 monitor's filters were retrieved at 5, 6, and 7 days later. Figure 15 provides a graph of the sample concentrations of each monitor for the run dates. Similar to the results of the Filter Extension Study, these results showed no significant change in concentrations from a filter retrieval time of 1 through 7 days.

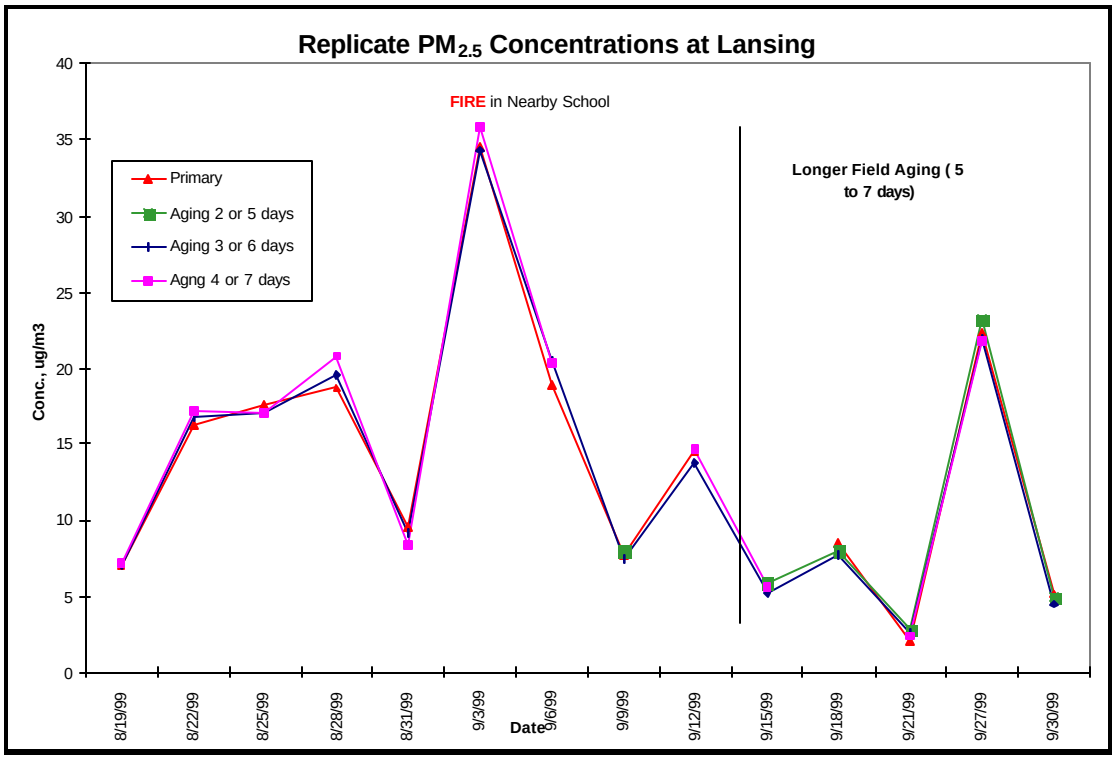
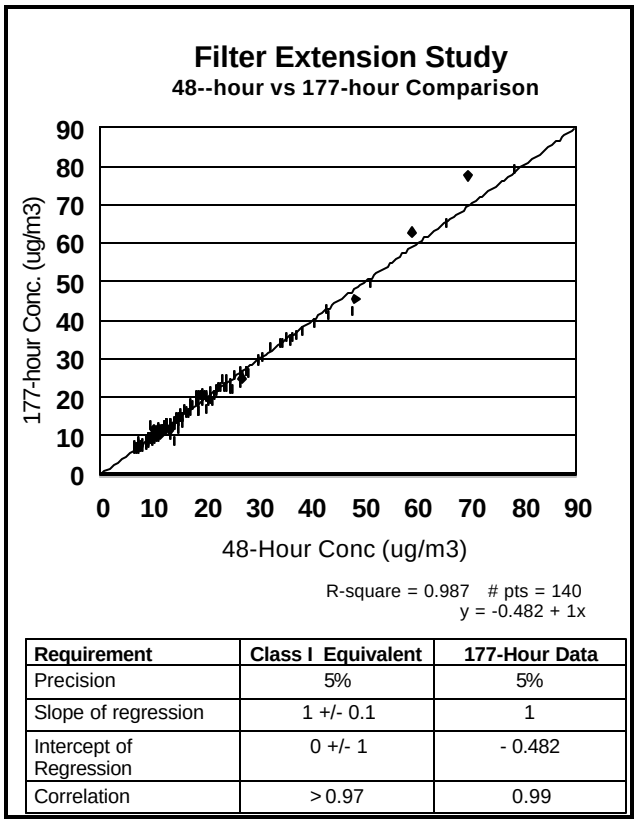


Figure 15. MI Study sample concentration results by sample date.

Conclusions

The filter extension study used the $PM_{2.5}$ data quality objectives for determining whether increasing the filter retrieval recovery period from 4 days (96 hours) to 7 days would increase measurement uncertainty to a level greater than the DQOs. All study sites average bias met the DQOs with a mean bias percent difference of -2.2% with 95% confidence intervals of -0.8% to -3.5%. The study results demonstrate that increasing this filter retrieval period did not compromise data quality at the study sites representing the ambient air monitoring program. As mentioned earlier in this document, even with a relaxation of this requirement, most samples in a sequential sampler would be recovered within the current 96 hour requirement regardless of the extension time allowance. Figure 16 provides one additional evaluation that is related to the Code of Federal Regulations for establishing Class I Federal Reference Method (FRM) equivalency. Although the Filter Extension study was not set up for establishing the 177-hour extension as a Class I equivalent method, in general, the test meets most of the requirements needed for the process. As the table in Figure 16 suggests, the 177-hour filter retrieval methods not only meet the $PM_{2.5}$ DQO criteria, but also meet the more restrictive Class I equivalent criteria, using the 48-hour data as the "reference method".



Requirement	Class I Equivalent	177-Hour Data
Precision	5%	5%
Slope of regression	1 +/- 0.1	1
Intercept of Regression	0 +/- 1	- 0.482
Correlation	> 0.97	0.99

Figure 16. Filter Extension Study comparison to CFR Federal Reference Method criteria

REFERENCES

USEPA (1997). Federal Register, Environmental Protection Agency, 40 CFR Parts 50 and 58. July 18, 1997.

USEPA (1998). Quality Assurance Guidance Document 2.12. Monitoring $PM_{2.5}$ in the Ambient Air Using Designated Reference or Class I Equivalent Methods. Environmental Protection Agency. November 1998.

USEPA (1999). Quality Assurance Report. Calendar Year 1999. The $PM_{2.5}$ Ambient Air Monitoring Program. Environmental Protection Agency. December 2000.

USEPA (2000). Quality Assurance Report. Calendar Year 2000. The SLAMS $PM_{2.5}$ Ambient Air Monitoring Program. Environmental Protection Agency. November 2001.

USEPA (1997). Data Quality Objectives for the $PM_{2.5}$ FRM program. Environmental Protection Agency. 1997