



UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

August 30, 2004

EPA-SAB-CASAC-CON-04-005

Honorable Michael O. Leavitt
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Subject: Clean Air Scientific Advisory Committee (CASAC) Ambient Air Monitoring and Methods (AAMM) Subcommittee Consultation on Methods for Measuring Coarse-Fraction Particulate Matter (PM_c) in Ambient Air (July 2004)

Dear Administrator Leavitt:

The Ambient Air Monitoring and Methods (AAMM) Subcommittee of the Clean Air Scientific Advisory Committee (CASAC) met in a public meeting held in Research Triangle Park (RTP), NC, on July 22, 2004, to conduct a consultation on methods for measuring coarse-fraction particulate matter (PM_c) in ambient air, based upon performance evaluation field studies conducted by EPA. Measurement of PM_c focuses on those particles in the ambient air with a nominal diameter in the range of 2.5 to 10 micrometers (*i.e.*, the coarse fraction of PM₁₀).

This project was requested by OAQPS in anticipation of the potential need for reference and equivalent methods for PM_c measurement, should new PM_c standards be established as a result of EPA's ongoing review of the national ambient air quality standards (NAAQS) for particulate matter (PM). The results of this consultation will support discussion of PM_c air quality monitoring to be included in the next draft of the OAQPS Staff Paper for PM, a policy assessment of scientific and technical information prepared as part of the PM NAAQS review. This draft Staff Paper is now planned for review by the CASAC PM Review Panel in early 2005.

As is our customary practice with a consultation, there will be no formal report from the CASAC or the SAB, nor does the Subcommittee expect any formal response from the Agency. Nevertheless, Subcommittee members were generally quite pleased with the effort that EPA has put into the development and field evaluation studies of coarse particle monitoring methods. The Subcommittee would welcome the opportunity to conduct a review of proposed PM_c reference and equivalent measurement methods in the future.

The roster of the CASAC AAMM Subcommittee is attached as Appendix A to this letter, and Subcommittee members' individual review comments are provided as Appendix B.

Sincerely,

/Signed/

Dr. Philip K. Hopke, Chair
Clean Air Scientific Advisory Committee

Appendix A – Roster of the CASAC AAMM Subcommittee

Appendix B – Review Comments from Individual CASAC AAMM Subcommittee Members

cc: Steve Page (MD-10)
Rich Scheffe (MD-14)
Tim Hanley (MD-14)
Michael Papp (MD-14)
Mary Ross (MD-15)

Robert Vanderpool (E205-01)
Thomas Ellestad (E205-01)
Anthony Maciorowski (1400F)
Fred Butterfield (1400F)

Appendix A – Roster of the CASAC AAMM Subcommittee

**U.S. Environmental Protection Agency
Science Advisory Board (SAB) Staff Office
Clean Air Scientific Advisory Committee
CASAC Ambient Air Monitoring and Methods (AAMM) Subcommittee***

CHAIR

Dr. Philip Hopke, Bayard D. Clarkson Distinguished Professor, Department of Chemical Engineering, Clarkson University, Potsdam, NY

Also Member: SAB Board

CASAC MEMBERS

Dr. Ellis Cowling, University Distinguished Professor At-Large, North Carolina State University, Colleges of Natural Resources and Agriculture and Life Sciences, North Carolina State University, Raleigh, NC

Mr. Richard L. Poirot, Environmental Analyst, Air Pollution Control Division, Department of Environmental Conservation, Vermont Agency of Natural Resources, Waterbury, VT

CONSULTANTS

Mr. George Allen, Senior Scientist, Northeast States for Coordinated Air Use Management (NESCAUM), Boston, MA

Dr. Judith Chow, Research Professor, Desert Research Institute, Air Resources Laboratory, University of Nevada, Reno, NV

Mr. Bart Croes, Chief, Research Division, California Air Resources Board, Sacramento, CA

Dr. Kenneth Demerjian, Professor and Director, Atmospheric Sciences Research Center, State University of New York, Albany, NY

Dr. Delbert Eatough, Professor of Chemistry, Chemistry and Biochemistry Department, Brigham Young University, Provo, UT

Mr. Eric Edgerton, President, Atmospheric Research & Analysis, Inc., Cary, NC

Mr. Henry (Dirk) Felton, Research Scientist, Division of Air Resources, Bureau of Air Quality Surveillance, New York State Department of Environmental Conservation, Albany, NY

Dr. Rudolf Husar, Professor, Mechanical Engineering, Engineering and Applied Science, Washington University, St. Louis, MO

Dr. Kazuhiko Ito, Assistant Professor, Environmental Medicine, School of Medicine, New York University, Tuxedo, NY

Dr. Donna Kenski, Data Analyst, Lake Michigan Air Directors Consortium, Des Plaines, IL

Dr. Thomas Lumley, Associate Professor, Biostatistics, School of Public Health and Community Medicine, University of Washington, Seattle, WA

Dr. Peter McMurry, Professor and Head, Department of Mechanical Engineering, Institute of Technology, University of Minnesota, Minneapolis, MN

Dr. Kimberly Prather, Professor, Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA

Dr. Armistead (Ted) Russell, Georgia Power Distinguished Professor of Environmental Engineering, Environmental Engineering Group, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA

Dr. Jay Turner, Associate Professor, Chemical Engineering Department, School of Engineering, Washington University, St. Louis, MO

Dr. Warren H. White, Visiting Professor, Crocker Nuclear Laboratory, University of California - Davis, Davis, CA

Dr. Yousheng Zeng, Air Quality Services Director, Providence Engineering & Environmental Group LLC, Baton Rouge, LA

SCIENCE ADVISORY BOARD STAFF

Mr. Fred Butterfield, CASAC Designated Federal Officer, 1200 Pennsylvania Avenue, N.W., Washington, DC, 20460, Phone: 202-343-9994, Fax: 202-233-0643 (butterfield.fred@epa.gov) (Physical/Courier/FedEx Address: Fred A. Butterfield, III, EPA Science Advisory Board Staff Office (Mail Code 1400F), Woodies Building, 1025 F Street, N.W., Room 3604, Washington, DC 20004, Telephone: 202-343-9994)

* Members of this CASAC Subcommittee consist of:

a. CASAC Members: Experts appointed to the statutory Clean Air Scientific Advisory Committee by the EPA Administrator; and

b. CASAC Consultants: Experts appointed by the SAB Staff Director to serve on one of the CASAC's standing subcommittees.

Appendix B – Review Comments from Individual CASAC AAMM Subcommittee Members

This appendix contains the preliminary and/or final written review comments of the individual members of the Clean Air Scientific Advisory Committee (CASAC) Ambient Air Monitoring and Methods (AAMM) Subcommittee who submitted such comments electronically. These comments are included here to provide both a full perspective and a range of individual views expressed by Subcommittee members during the Subcommittee's Consultation on Methods for Measuring Coarse-Fraction Particulate Matter (PM_c) in Ambient Air. These comments do not represent the consensus views of the CASAC AAMM Subcommittee, the CASAC, the EPA Science Advisory Board, or the EPA itself. The list of Subcommittee members providing comments is provided on the next page, and their individual comments follow.

<u>Panelist</u>	<u>Page #</u>
Dr. Philip Hopke	B-3
Dr. Ellis Cowling	B-4
Mr. Richard L. Poirot	B-7
Mr. George Allen	B-18
Dr. Judith Chow	B-23
Mr. Bart Croes	B-33
Dr. Kenneth L. Demerjian	B-37
Dr. Delbert Eatough	B-40
Mr. Eric Edgerton	B-46
Mr. Henry (Dirk) Felton	B-48
Dr. Rudolf Husar	B-54
Dr. Kazuhiko Ito	B-57
Dr. Donna Kenski	B-62
Dr. Thomas Lumley	B-65
Dr. Peter McMurry	B-67
Dr. Kimberly Prather	B-71
Dr. Armistead (Ted) Russell.....	B-74
Dr. Jay Turner	B-77
Dr. Warren H. White	B-81
Dr. Yousheng Zeng	B-84

Dr. Philip Hopke

Comments on the Development of Monitors in Support of a NAAQS for PM_(10-2.5)

Philip K. Hopke
Bayard D. Clarkson Distinguished Professor and
Director, Center for Air Resources Engineering and Science
Clarkson University

In the development of monitoring methods for PM_(10-2.5), there are two major objectives to be achieved: measurement of the attainment/non-attainment status of a particular location and the determination of the composition of the particles in support of air quality planning in the event of non-attainment of the standard. The conventional approach has been to deploy filter-based samplers on the premise that the samples could be used for subsequent chemical analysis. However, such an approach only produces at best, 1 24-hour integrated value every third day. In general these samples have not been extensively used for chemical characterization and the cost of collecting the samples given the amount of labor in preparing, deploying, and weighing the exposed filters is very high for the very limited amount of data produced. Thus, the primary focus of any deployment of monitors in support of a coarse particle NAAQS should be on continuous monitors. It was clear from the presentations that such monitors are sufficiently close to being ready for deployment that they can be considered for FRM status, particularly those that are based on fundamental physical principles. Given the limited amount of PM_(10-2.5) data and epidemiology available on which to base a standard, it would be better to start with a continuous monitor and have sufficient data to permit additional health effects epidemiology to be performed so that refined standards could be promulgated during the next round of review. Filter samplers should only be deployed in areas of likely non-attainment so that the compositional information would be available to support of air quality planning.

I would strongly recommend a performance standard rather than a design standard. The lack of precision in the 1987 PM₁₀ FRM was due to a lack of rigor in the specifications of the sampler requirements. There was no requirement for flow control and the performance envelop on the inlet was set too wide. From simulations based on measured size distributions of the ambient atmosphere, it is possible to develop an inlet performance envelop that would provide the precision necessary to meet defined data quality objectives. Thus, the key would be to define the level of precision required to permit effective attainment/non-attainment decisions to be made and then develop a set of performance criteria that are needed to achieve this level of precision. These criteria may prove a challenge for the instrument manufacturers, but will not produce the substantial impediment to technological innovation that design standards represent.

Dr. Ellis Cowling

[Note: Sent via e-mail to CASAC Chair Dr. Philip Hopke, members of the CASAC Particulate Matter (PM) Review Panel, and the CASAC Designated Federal Officer (DFO) at 12:22 PM on August 2, 2004]

In general, I find substantial merit in this First Draft summary of CASAC comments. I believe these comments provide valuable guidance for NCEA's further efforts to provide a draft Air Quality Criteria Document for Particulate Matter that can be accepted in full by CASAC at its forthcoming conference call discussion — now tentatively scheduled for some time in mid September.

As befits my special particular role in CASAC, my major concerns about the AQCD for PM have to do with the need for a more balanced treatment in the AQCD for PM of “Welfare Effects,” and the associated desirability of a “Secondary Standard” dealing with PM effects on various “Air-Quality Related Values.”

These values include: 1) visibility impacts on human enjoyment of scenic vistas especially in national and state parks, 2) associated economic impacts on our tourism industries, 3) ecosystem responses to decreased solar radiation caused by regional haze, 4) increased atmospheric deposition of the nutrient and growth-altering substances in PM (including organic, oxidized, and reduced forms of nitrogen, sulfur, phosphorus, potassium, and the wide variety organic nutrients of fine and coarse aerosol particles, 5) direct effects on materials such as soiling of painted surfaces, exposed textile materials, etc., and 6) the need for greater concern during the next several decades about “smoke management” in light of the greatly increased risk of wild fires and the corresponding necessity for increased amounts of controlled burning of forests and natural areas in parks and other recreational areas.

Greater attention should be given in the AQCD to these “Air Quality Related Values” in rural as well as in urban areas.

Some of the many excellent and readily available photographs, tables, and figures should be added to the AQCD to illustrate and quantify such welfare effects as: 1) visibility impairment at scenic vistas and airports, 2) wild fire impacts on the aesthetic values of landscapes, 3) wildfire impacts on wildlife populations, 4) economic data on tourism impacts and smoke management costs and benefits, 5) the successes of urban areas that have adopted secondary standards for visibility impairment, and 6) changes in populations of aquatic invertebrates or fish that are induced by atmospheric deposition of the essential nutrient substances in the aerosols involved in cloud nucleation and precipitation processes.

With regard to ideas for inclusion in the summary of individual comments deriving from the CASC “Consultation on Methods for Measuring Coarse-Fraction Particulate Matter (PM_c) in Ambient Air, Based upon Performance Evaluation Studies Conducted by EPA,” permit me to summarize the two points I made at near the end of this “Consultation” on Thursday July 22.

Point 1)

EPA (and many other federal research and monitoring organizations) need to guard against the tendency to allocate so much of the funds used in field measurement campaigns to “making careful measurements” and that inadequate funds are available for “scientific analysis and interpretation” to determine what the measurements really mean.

As described on pages 282-284 in the attachment to this E-mail message, these cautionary remarks about problems in field measurement programs were suggested originally by the late Glenn Cass, formerly of Cal Tech and later of Georgia Tech, on the basis of his career-long experience in various environmental monitoring programs — programs in which too much funding was allocated to “measurements” and too little to “analysis and interpretation” of the data. Please note on pages 283 and 284, the “Fifteen general and specific reasons why this happens” and the “Thirteen general and specific things that can be done about it!”

The reference for this published reviewed paper is: Cowling, E., and J. Nilsson. 1995. Acidification Research: Lessons from History and Visions of Environmental Futures. *Water Air and Soil Pollution* 85:279-292.

Please also note especially the suggestion in item 9 on page 284 about a “50:50 distribution” of funding allocations between “measurements” and “analysis and interpretation” of monitoring data rather than the (90:10 or 80:20 distribution) that is typical of many monitoring programs in EPA and other agencies.

But please also note that an even better suggestion was made by Mary Barber, former executive leader with the Ecological Society of America’s Executive, who opposed the “50:50 distribution” idea at a recent Whitehouse Conference on monitoring. Mary Barber insisted, and I agree with her, that it would be even more appropriate to distribute the funding into three rather than two categories of investments — with equal shares going to “measurements,” “analysis and interpretation,” and “outreach and extension of findings” to interested clientele and “customers” for the results of field measurement programs.

This problem is so commonplace — not only in this country but all over the world — that I commend these “lessons that are available to be learned” (and perhaps even the “15 reasons why this happens” and the “13 things to do about it”) for inclusion among the “comments from individual participants” in the CASAC Consultation on PM Measurement Methods.

Point 2)

EPA should also guard against the tendency to give undue emphasis to “Data Quality Objectives” in the selection and evaluation of instruments and subsequent implementation of field measurement programs to the exclusion of concern about “Science Quality Objectives” and “Policy Relevancy Objectives.”

Experience within the Southern Oxidants Study and other large-scale field measurement campaigns have demonstrated repeatedly that undue emphasis on “Data Quality Objectives” often leads to:

- 1) Serious lack of attention to the scientific hypotheses and assumptions that are inherent in any choice of scientific instruments, the appropriateness of the ground-based sites or aircraft platforms on which the instruments are mounted, the skills of the instrument operators, the data processing and data-display programs used, and especially the scientific quality of the conclusions and statements of findings that are drawn from analysis and interpretation of the measurements that are made; and

2) Equally serious lack of attention to the policy relevancy of the measurements being made — relevancy to the general or specific enhancements of environmental protection that are the *raison de etre* of the public health or public welfare concerns that led to the decision to establish a monitoring program or undertake a field measurements campaign in the first place.

In this latter connection, permit me also to call attention to the “Guidelines for the Formulation of Scientific Findings to be Used for Policy Purposes” which were developed originally by the NAPAP Oversight Review Board led by Milton Russell. Please find attached to this E-mail message, an electronic version of these Guidelines which we have adopted and very slightly adapted for use in formulating policy relevant scientific findings in the Southern Oxidants Study.

The original version of these Guidelines was published as Appendix III of the April 1999 Report titled “The Experience and Legacy of NAPAP.” This was a Report to the Joint Chairs Council of the Interagency Task Force on Acidic Deposition of the Oversight Review Board (ORB) of the National Acid Precipitation Assessment Program.

As indicated in Appendix III:

“The following guidelines in the form of checklist questions were developed by the ORB to assist scientists in formulating presentations of research results to be used in policy decision processes. These guidelines may have broader utility in other programs at the interface of science and public policy and are presented here with that potential use in mind.”

These guidelines may also be of value as part of the “communication of individual comments” from the CASAC Consultation on PM Measurement Methods.

Dr. Ellis B. Cowling, Director
Southern Oxidants Study
North Carolina State University
Raleigh, North

Mr. Richard L. Poirot

Supplemental Comments on EPA PM_c Methods Review, R. Poirot, 8/1/04

Overall, the effort EPA (& equipment manufacturers) have put into development and field evaluation of coarse particle methods is excellent, and should be commended and continued (with adequate funding support). Given the substantial expertise of many of the subcommittee members on the technical details of the sampling methods, my comments are intentionally of a more general nature, and specifically encourage EPA to consider the many important objectives of a new PM_c sampling program – in addition to determining compliance with anticipated new standards.

Towards the beginning of the 7/22/04 CASAC AAMM Subcommittee meeting, I raised some general (devil's advocate) "issues" relating to the multiple objectives for PM_c monitoring, with the intent of generating some group discussion. In hindsight, this didn't work out so well – and I'd like to try to clarify some of these points.

The draft chapters 8 & 9 of the PM CD seem to (me to) strongly and repeatedly emphasize the relatively benign nature of "crustal" PM. Since crustal material (i.e. soil and its associated physical, chemical & biological components) typically account for a large proportion of coarse particle mass, this might raise questions on the focus of coarse mass as a relevant health effects indicator. I don't agree with this emphasis in the current draft CD chapters, think that it is overstated in a speculative manner, and have submitted supplemental comments on this issue (copy attached at bottom of these NAAMM comments).

Those health studies that have specifically attempted to evaluate effects of coarse particles have shown a wide range of potential responses – ranging from (illogical) negative associations with mortality, to large positive associations that exceed those for fine particles. Although it is currently unclear what level(s) or averaging time(s) will be proposed for a PM_c standard(s), a reasonable assumption is that there will not be high confidence that any specific level of standard will represent a clear "bright line" above and below which effects do and don't occur. Given the anticipated high degree of spatial and temporal variability of ambient PM_c concentrations (compared to PM_{2.5}), it also seems logical that we will have less confidence in how well any new PM_c measurements represent human exposures. For these reasons, I suggest that the (always important) monitoring objective of determining compliance with standards is of relatively less importance for PM_c than it is for many other criteria pollutants (PM_{2.5} & ozone for example). Conversely, other measurement objectives take on relatively greater importance and should be carefully considered.

Virtually all past studies that have jointly considered health effects of fine and coarse particles have employed methods that are more precise for PM_{2.5} than for PM_c – with the latter typically quantified by subtracting (precise) PM_{2.5} from imprecise PM₁₀ (and with the subtraction further compounding the noise of the individual measurements). As emphasized in Warren White's comments & associated reference (White, 1998, *J. Air & Waste Manage. Assoc.* 48, 454-458): "noise always depresses observed correlations with other measurements, historical PM_c

measurements have been much noisier than associated PM_{2.5} measurements, and crucial inferences about health effects and atmospheric behavior are routinely based on the differences observed between correlations involving PM_c and PM_{2.5}.” Thus the current uncertainty over specific PM_c effects is at least partially due to the absence of collocated PM_c and PM_{2.5} of comparable precision (and, for epidemiological studies) in locations representative of large population exposures. For these reasons, an especially important objective of new PM_c measurements should be to collocate precise PM_c and PM_{2.5} both in areas of high anticipated PM_c exposures (arid southwest and agricultural areas), and also in large population centers nationwide. Given the anticipated high degree of temporal variability in PM_c (vs. PM_{2.5}), the development of precise continuous methods for PM_c seems especially important for meeting this objective, in support of future health studies.

For the most part, the chemical composition of coarse particles has not, historically, been well characterized (typically done by subtraction or dicot if at all) and this would seem like a third important objective. Here, I think the sampling separation of PM_c from PM_{2.5} deserves special emphasis. One of the original design concepts of the dichotomous sampler was to prevent the mixing (and subsequent chemical interaction) of fine & coarse particles (for example to quantify fine particle acidity without artifactual reaction with alkaline crustal material - & vice versa). For this reason, methods that allow collection of coarse-only particles on filter media suitable for chemical analyses should also be emphasized. For this reason, I encourage EPA to consider reanalysis (by other labs and methods) of the coarse dicot filters (or maybe the PM-10 and 2.5 filters) collected during the field evaluations – with a specific emphasis on developing closure with the measured mass. Also, some evaluation of the uniformity of coarse particle distribution across filter media and/or any limitations of surface beam (XRF) analytical methods on coarse particle samples should be evaluated – possibly through comparison with neutron activation analysis.

Along similar lines, one suggestion for PM_c methods development is to consider – or at least evaluate the effects of - adding SO₂ & HNO₃ denuders to PM_c samplers to prevent artifactual reactions with alkaline crustal or marine sea salt material during sampling. Potentially, the use of such denuders would be most important for 24-hour filter-based methods and less critical for “fast response” continuous methods.

Given the above-stated current uncertainties in absolute “bright line” PM_c concentrations which are injurious, the high temporal variability, and the greater artifact potential, I don’t think it’s absolutely necessary (and perhaps not desirable) to have a filter-based FRM for PM_c. Some consideration should be given to specifying a continuous method as the FRM, with a filter method as FEM, specifically intended for subsequent speciation analyses at a limited number of sites.

For various reasons, developing a (much) better understanding of the biological content of coarse particles seems especially important. It’s unclear to me if there are currently available analyses that could reveal key biological indicators from filter-based samples on a “routine” basis, but I would suggest EPA devote some research into potential, low-cost “biological indicator” analyses that might be conducted on “routine” PM_c filter samples (protein, carbohydrates, etc.), and also consider development of PM_c sample collection methods that would allow a more

comprehensive evaluation of specific biological components (pollen, endotoxin, and specific fungi, virus and/or bacteria). Clearly such detailed biological analyses will not be economically feasible on a large network basis, but it would be important to devote some fraction of monitoring (&/or research) funds to development & implementation of such methods at a limited number of sites.

It can be reasonably assumed in advance that PM_c will exhibit a high degree of spatial variability (especially for locally generated particles – PM_c from more distant sources like African or Asian dust storms will be more spatially uniform). For this reason, I caution against attempting to quantify this spatial variability by uniformly dense siting of routine monitors (we can't afford it), and encourage EPA to consider alternative methods of characterizing spatial variability of concentrations & exposures. Characterizing PM_c spatial patterns can't be done by measurements alone, and development of better microscale models should be considered an important objective. Characterizing temporal variability with continuous methods will help here, as time/space patterns are related. Better quantification of the full range of coarse particle size distributions (including above 10 μm) such as by optical particle counters will also be helpful. Careful attention to monitor siting characteristics, especially inlet height and distance from roadways, agricultural and mining activities will also be critical. Focused gradient studies (how rapidly do concentrations fall off with distance from roadways?) will be useful (an important co-objective of such studies should be to clarify the current, continuing disparity between EPA PM_{2.5} road dust emissions estimates and ambient concentrations!). But such “experiments” do not need to be repeated everywhere, and consideration should be given to identifying a small number of more intensive sites in each region of the country. Possibly such sites could be periodically “rotated” among states in a region to avoid disruption of state-specific monitoring funds.

While filter-based measurements of PM_{2.5} and low-volume PM₁₀ with subsequent subtraction is the least desirable measurement approach, it should be recognized that low volume sampling – in many cases collocated with PM_{2.5} and/or PM_{2.5} speciation samplers is or will be conducted for “air toxics” purposes by State agencies. Some consideration should be given to coordination between the toxics and PM_c programs, to at least allow for use of subtraction (of precise samples) at sites where such measurements are being conducted already. In addition, for sites where PM_{2.5} speciation measurements are collocated with low-vol PM₁₀, some consideration might be given to providing for complementary XRF analysis on the PM₁₀ samples (I think ICPMS is generally the most common method for “toxics metals”, but possibly it could be preceded by (nondestructive) XRF at selected sites.

Along similar lines, is there any possibility of modifying inlets of existing PM_{2.5} speciation samplers to a 10 μm cut? My sense is the PM_{2.5} speciation samplers are more widely deployed than need be, and if a few of these could be converted to PM₁₀ and collocated at a few PM_{2.5} speciation sites, we could possibly generate some comparative data relatively rapidly and at modest cost.

[The following is also included at the request of Mr. Poirot:]

Supplemental Comments on PM CD Chapters 8 and 9, R. Poirot 7/26/04

Several sections of chapters 8 & 9 (for example 8.2.2.5.3 & 9.2.3.2.1) summarize health effects associations with different chemical components and/or source categories on PM in various size fractions. These discussions are clear, detailed, helpful and informative – and (I think) the results could conceptually be presented, integrated, summarized, etc. in two general ways:

1. to indicate that many or most all of the major PM mass-contributing species or source contributions have been individually shown to be injurious (this adds considerably to the use of PM_{2.5} or PM_{10-2.5} mass as a regulatory metrics, regardless of different PM mixtures in different regions).
2. to indicate that some species or source categories appear to be more harmful or less harmful than others (potentially this might lead to species-specific standards or source-specific priorities in the implementation phase).

Based on discussions at the CASAC PM CD review, subsequent discussions at the PM coarse monitoring methods review and on a re-reading of relevant sections of chapters 8 and 9, I encourage EPA to more heavily emphasize the former (#1) use of this information and de-emphasize the latter (#2). General reasons include:

- a. Adverse health effects are associated with many different species and/or source-specific contributions, although these associations are not always consistent among studies. Taken in the aggregate, they clearly show adverse effects from many species, but individually no one study is definitive.
- b. The species and/or source-specific health associations are not sufficiently strong or consistent in their findings to support species-specific standards or to prioritize (or exculpate) species or sources for future controls at the present time – and to do so would require choosing among or rating studies which show contrary effects (a much more difficult argument to support than #1).
- c. Epi. studies associating specific source categories with effects (or non-effects) are limited in number, and have generally have relied on “factor analysis” approaches (such as PCA with Varimax or Procrustean rotation) which are not currently considered state-of-the-science (poorly constrained and potentially yielding many different “equally correct” answers) and require subjective interpretation of the resulting sources. These results are then often further interpreted and commented on in the CD in a highly speculative manner.

Specifically, I think the chapter 9 integrated synthesis should de-emphasize or present counter examples in sections where specific source categories are identified as uniquely benign. This seems most evident for the contributions of “crustal” emissions to PM_{2.5}, PM₁₀ and/or PM_{10-2.5}. I think this is especially not helpful in considering any coarse particle standards, since

crustal material (and its associated anthropogenic chemical or biological contaminants) is typically a large fraction of coarse mass at most times & places. For example following are several examples where I think the potential effects of “crustal particles” are unnecessarily (& speculatively) de-emphasized:

On p. 9-44, lines 18-19: “Also of much importance, all of the above studies that investigated multiple source categories found a soil or crustal source that was negatively associated with mortality”. Here, it’s not entirely clear why this is “of much importance” (compared to what?), or what “all of the above studies” refers to (the preceding paragraph, page, section?). The consistent finding of a negative association (and implication we would live longer if it were dustier) is a consistent indication (to me) of a poorly formulated model(s). It is also inconsistent with the many studies (mostly cited in the CD) which do show effects associated with coarse particle mass, and with the rather extensive bodies of literature on adverse effects from both the inorganic components of crustal material (silicosis, pneumoconiosis, etc.), as well as with the extensive and growing literature on diseases associated with soil-borne fungi or bacteria (Coccidioidomycosis, etc.). I’ve listed some references grouped in these 3 general areas at the end of these comments.

Several features of the (rather outdated) receptor model approach taken by the studies which I assume are referred to in “all of the above studies” are important. First, all multi-elemental measurement techniques, and especially the most common XRF, coincidentally quantify a large number of elements which are of predominantly crustal origin (Si, Al, Fe, Ca, Ti, etc. – much more so than for any other source category). For this reason, a “crustal” or “soil” factor is nearly always identified in virtually all receptor model applications. The (rotated eigenvector) factor analysis approach which I think was used in all of the above studies seeks first to account for the collective variance of all the species used as input, and so typically (prior to rotation) the first component, explaining a maximum of the total variance tends to be “crustal” (even though these elements together typically account for only a small fraction of the fine mass). Subsequent rotational schemes (Varimax, Procrustes, etc.) then redistribute the variance in ways that require highly subjective decisions by the modelers. These models also require (can only find) sources of fixed, unique chemical composition and variable, unique contribution. Soil itself has a highly variable composition but tends to be more alkaline in the West than in the East, very alkaline in areas with calcareous bedrock, and different yet again in the Sahara Dust and Asian Dust which often result in the highest soil contributions in the Eastern US and West coasts respectively. These more distant dust events also tend to have much smaller particle size distributions than “local dust” emissions, as the larger particles are more readily removed during transport. Crustal material can become heavily contaminated with anthropogenic S, N, OC, EC, salt and metals - both as it is deposited & resuspended from roadways or as it undergoes chemical reactions during transport. Conversely, many other sources also contain “crustal impurities” (coal fly ash for example), and so when one obtains a “pure crustal source” from a factor analysis it’s not entirely clear what that source actually represents. If the rotation is oblique, the sources are required to be uncorrelated, and it’s therefore highly probable that the “crustal” source will (to the extent local sources contribute) be a good indicator of high wind speeds, since this will lead uniquely to high emissions & concentrations of dust which will be uncorrelated with all other (gaseous &) particulate pollutants. While high dust concentrations that also build up under stagnation conditions (from road dust emissions) or dust from more distant origins will tend to

get mixed into other modeled sources. Quite possibly the consistent finding of negative health associations with dust just reflects windy days when folks stay indoors and the air is otherwise at its cleanest. For example:

On p. 9-27, lines 1&2, we learn that “new studies have shown no increases in mortality on days with high concentrations of wind-blown dust (crustal particles), using PM 10 concentrations and data on wind speed as indicators of dust storm days.” Which new studies? I think the (not unreasonable) use of wind speed as a dust surrogate is telling, as dust emissions (especially the maximum concentrations) are uniquely associated with high wind speeds – which in turn will tend to minimize concentrations of all other (fine) particle and gaseous components – assuring minimal chemical reactions between crustal particles and other species. High concentrations of crustal particles and chemically associated contaminants (on the surface of coarse particles) from MV, SO₂ or smelting activities would also reach high concentrations (as would many other gaseous and PM pollutants) on local stagnation days with low mixing heights – but would not be considered with this “wind speed” surrogate (nor would dust of distant origin). Potentially outdoor activities are curtailed on very windy, “local” dusty days, windows are closed, inhalation efficiency of coarse particles likely decreases with wind speed, and the spatial representativeness of “central site monitors” diminishes. Conversely, the lengthy Section 8.4.3.5 discussion of “Adjustments for Meteorological Variables” includes factors like temperature and humidity that might tend to exaggerate assumed PM effects, but makes no mention of wind speed - which might tend to diminish such effects.

On p. 9-27, lines 3-6, it is postulated that cardiovascular mortality in Phoenix may be due to the metal rather than crustal content of coarse particles. Yet on p. 8-63, lines 22-28 it’s indicated that “... (Smith et al., 2000) indicate that coarse particle-mortality associations are stronger in spring and summer, when the anthropogenic metal (Fe, Cu, Zn, and Pb) contribution to PM_{10-2.5} is lowest, as determined by factor analysis.” In this case, the seasonal association of effects when crustal, not metal, coarse particles are greatest is attributed speculatively to “biogenic processes (e.g., wind-blown pollen fragments, fungal materials, endotoxins, and glucans) of the particles during spring and summer”. It is also specifically emphasized that the authors “observed that the implication that crustal, rather than anthropogenic elements, for the observed relationship with mortality was counterintuitive.” Thus a finding that does not fit the theory is discredited.

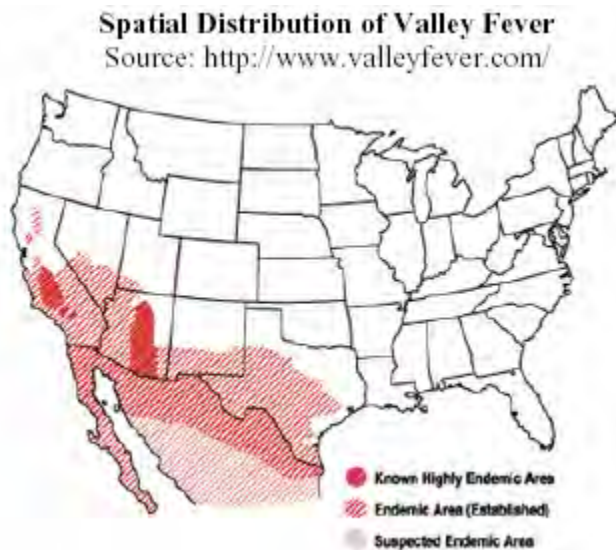
Emphasizing the potential importance of coarse biological content is reasonable, but on p. 8-326, lines 8-17, its indicated that “Reasons for differences among findings on coarse-particle health effects reported for different cities are still poorly understood, but several of the locations where significant PM_{10-2.5} effects have been observed (e.g., Phoenix, Mexico City, Santiago) tend to be in drier climates and may have contributions to observed effects due to higher levels of organic particles from biogenic processes (e.g., endotoxins, fungi, etc.) during warm months.”

Here, I can understand how dry climate can and does lead to increased emissions and concentrations of coarse crustal material (and any biological material it contains), but I’m not sure why or if its logical to expect arid climates (and associated sparse vegetation) to have uniquely higher pollen, endotoxin or fungi emissions & concentrations than humid areas – where wind-blown dust emissions would tend to be suppressed by precipitation, and where pollens, pollen fragments and fungi might be relatively more abundant.

I think a more logical explanation could be effects from soil-associated fungi, which for the most part become airborne only as the soil becomes airborne during “natural” dust storms and/or as modified by human agricultural activities (tilling harvesting, grazing, etc.) and on & off-road vehicles.

For example, the geographically-focused incidence of “Valley Fever” specifically caused by the fungus *Coccidioides sp.*, which grows in soils in areas of low rainfall, high summer temperatures, moderate winter temperatures, and which is emitted in direct association with the soil that supports it, would seem like a more logical causal or contributing factor than some non-soil-related biogenic contribution from pollen or more benign fungi in general. See also the references on other soil-related fungal or bacteriological effects on human & animal health, crops, aquatic ecosystems, etc. – for example Garrison et al. (2003).

On p. 8-326, lines 17-21 it is indicated that “in some U.S. cities (especially in the NW and the SW) where PM_{10-2.5} tends to be a large fraction of PM₁₀, measurements, coarse thoracic particles from wood burning are often an important source during at least some seasons. In such situations, the relationship between hospital admissions and PM₁₀ may be an indicator of response to coarse thoracic particles from wood burning.”



However, since wood smoke concentrations are VERY predominately $< 2.5 \mu\text{m}$, it seems illogical that wood smoke should be the likely causal factor for coarse particle effects in areas that have high coarse: fine ratios. I also question whether the NW has a high coarse: fine ratio and why the (dusty, crusty) SW would tend to have a uniquely high coarse wood smoke contribution (compared to all northern areas where space heating demands and fuel wood supplies are greater). This also seems inconsistent with the “counterintuitive” Phoenix results indicating highest coarse PM effects in the spring & summer. I’m getting picky here, but again it looks like trying too hard to show “it must be anything but crustal emissions”...

References

On Coarse PM health effects associations in general:

Becker S, Soukup J. Coarse(PM(2.5-10)), fine(PM(2.5)), and ultrafine air pollution particles induce/increase immune costimulatory receptors on human blood-derived monocytes but not on alveolar macrophages. *J Toxicol Environ Health A*. 2003 May 9;66(9):847-59.

Becker S, Soukup JM, Sioutas C, Cassee FR. Response of human alveolar macrophages to ultrafine, fine, and coarse urban air pollution particles. *Exp Lung Res*. 2003 Jan-Feb;29(1):29-44.

Becker S, Fenton MJ, Soukup JM. Involvement of microbial components and toll-like receptors 2 and 4 in cytokine responses to air pollution particles. *Am J Respir Cell Mol Biol*. 2002 Nov;27(5):611-8.

Soukup JM, Becker S. Human alveolar macrophage responses to air pollution particulates are associated with insoluble components of coarse material, including particulate endotoxin. *Toxicol Appl Pharmacol*. 2001 Feb 15;171:20-6.

Ostro BD, Broadwin R, Lipsett MJ. Coarse and fine particles and daily mortality in the Coachella Valley, California: a follow-up study. *J Expo Anal Environ Epidemiol*. 2000 Sep-Oct;10(5):412-9.

Ostro BD, Hurley S, Lipsett MJ. Air pollution and daily mortality in the Coachella Valley, California: a study of PM10 dominated by coarse particles. *Environ Res*. 1999 Oct;81(3):231-8

Sheppard L, Levy D, Norris G, Larson TV, Koenig JQ. Effects of ambient air pollution on nonelderly asthma hospital admissions in Seattle, Washington, 1987-1994. *Epidemiology*. 1999 Jan;10(1):23-30.

Cifuentes LA, Vega J, Kopfer K, Lave LB. Effect of the fine fraction of particulate matter versus the coarse mass and other pollutants on daily mortality in Santiago, Chile. *J Air Waste Manag Assoc*. 2000 Aug;50(8):1287-98. Burnett et al. 2000

Lin M, Chen Y, Burnett RT, Villeneuve PJ, Krewski D. The influence of ambient coarse particulate matter on asthma hospitalization in children: case-crossover and time-series analyses. *Environ Health Perspect*. 2002 Jun;110(6):575-81. Sheppard L et al. 1999

Pozzi R, De Berardis B, Paoletti L, Guastadisegni C. Inflammatory mediators induced by coarse (PM2.5-10) and fine (PM2.5) urban air particles in RAW 264.7 cells. *Toxicology*. 2003 Feb 1;183(1-3):243-54.

Kleinman MT, Sioutas C, Chang MC, Boere AJ, Cassee FR. Ambient fine and coarse particle suppression of alveolar macrophage functions. *Toxicol Lett*. 2003 Feb 3;137(3):151-8.

Monn C, Becker S. Cytotoxicity and induction of proinflammatory cytokines from human monocytes exposed to fine (PM2.5) and coarse particles (PM10-2.5) in outdoor and indoor air. *Toxicol Appl Pharmacol* 1999;155: 245-52.

Schins RP, Lightbody JH, Borm PJ, et al. Inflammatory effects of coarse and fine particulate matter in relation to chemical and biological constituents. *Toxicol Appl Pharmacol*. 2004;195(1): 1-11.

Shi T, Knaapen AM, Begerow J, et al. Temporal variation of hydroxyl radical generation and 8-hydroxy-2'-deoxyguanosine formation by coarse and fine particulate matter. *Occup Environ Med*. 2003;60: 315-21.

Greenwell LL, Moreno T, Jones TP, et al. Particle-induced oxidative damage is ameliorated by pulmonary antioxidants. *Free Rad Biol Med*. 2002;32(9): 898-905.

Mar TF, Norris GA, Koenig JQ, et al. Associations between air pollution and mortality in Phoenix, 1995-1997. *Environ Health Perspect* 2000;108: 347-53.

Castillejos M, Borja-Aburto V, Dockery D, et al. Airborne coarse particles and mortality. *Inhal Toxicol* 2000; 12 (supplement 1):61-72.

On dust –associated inorganic components & effects:

Gift, J.S., & Faust, R.A.: “Noncancer Inhalation Toxicology of Crystalline Silica: Exposure-Response Assessment,” *J. Expos. Analysis Environ. Epidemiol.* 7(3):345-358 (1997).

Wright, G.W.: “The Pulmonary Effects of Inhaled Inorganic Dust” (Chapter 7). In *Patty’s Industrial Hygiene and Toxicology*, 3rd Rev. Ed., Vol. 1 — General Principles. New York: John Wiley & Sons, 1978. pp. 175-176.

M. Schenker (2000) Exposures and Health Effects from Inorganic Agricultural Dusts, *Environmental Health Perspectives Supplements*, Volume 108, Number S4 August 2000.

International Agency for Research on Cancer: *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 68 — Silica, Some Silicates, Coal Dust and Paraaramid Fibrils*. Geneva, Switzerland: World Health Organization, IARC Press, 1997. pp. 41-85.

U.S. Environmental Protection Agency/Office of Research and Development: *Ambient Levels and Noncancer Health Effects of Inhaled Crystalline and Amorphous Silica: Health Issue Assessment* (EPA/600/R-95/115). November 1996. pp. 7-1–7-6.

On dust –associated biological components & effects:

Buchwaldt., L., ... and C.C. Bernier. 1996. Windborne dispersal of *Colletotrichum truncatum* and survival in infested lentil debris. *Ecology and Epidemiology* 86:1193. Gloster, J., R.F. Sellers, and A.I. Donaldson. 1982. Long distance transport of foot-and-mouth disease virus over

the sea. *The Veterinary Record* 110(Jan. 16):47. Griffin, D.W., V.H. Garrison, *et al.* 2001. African desert dust in the Caribbean atmosphere: Microbiology and public health. *Aerobiologia* 17(June 14):203.

Griffin, D.W., C.A. Kellogg, and E.A. Shinn. 2001. Dust in the wind: Long range transport of dust in the atmosphere and its implications for global public and ecosystem health. *Global Change and Human Health* (September). O'Hara, S.L., *et al.* 2000. Exposure to airborne dust contaminated with pesticide in the Aral Sea region. *Lancet* 355(Feb. 19):627.

Prospero, J.M. 1999. Long-term measurements of the transport of African mineral dust to the southeastern United States: Implications for regional air quality. *Journal of Geophysical Research* 104(July 20):15917.

Snyder LS, Galgiani JN. Coccidioidomycosis: The Initial Pulmonary Infection and Beyond. *Seminars in Respiratory and Critical Care Medicine* 1997;18:235-247.

Galgiani JN. Coccidioidomycosis. *In:* Remington JS and Swartz MN ed. *Current Clinical Topics in Infectious Diseases*. 1997, pp. 188-204.

Morbidity & Mortality Weekly Report, "Coccidioidomycosis- Arizona, 1990-1995. Dec. 13, 1996/vol. 45:1069-1073.

Einstein HE, Johnson RH. Coccidioidomycosis: new aspects of epidemiology and therapy. *Clinical Infectious Diseases* 1993;16:349-356.

Galgiani JN. Coccidioidomycosis. *Western Journal of Medicine* 1993;159:153-171.

Galgiani JN. Coccidioidomycosis. *In:* Rakel RE, ed. *Conn's Current Therapy*. Philadelphia: WB Saunders Co., 1996, pp. 188-190.

Galgiani JN, Ampel NM. Coccidioidomycosis in human immunodeficiency virus-infected patients. *Journal of Infectious Diseases* 1990;162:1165-69.

Hall KA, Copeland JG, Zukoski CF, *et al.* Markers of coccidioidomycosis before cardiac or renal transplantation and the risk of recurrent infection. *Transplantation* 1993;55:1422-24.

Kwon-Chung KJ, Bennett JE. *Medical Mycology*. Philadelphia, Lea and Febiger, 1992, pp. 356-396.

Pappagianis D. Marked increase in cases of coccidioidomycosis in California: 1991, 1992, and 1993. *Clinical Infectious Diseases* 1994;19(Suppl 1):S14-18

Pappagianis D, Zimmer BL. Serology of coccidioidomycosis. *Clinical Microbiology Reviews* 1990;3:247-268.

Schneider E, et al: A Coccidioidomycosis Outbreak Following the Northridge, Calif Earthquake. JAMA 1997;277:904-908.

Pappagianis D: Coccidioidomycosis, *In* DiSalvo A (Ed) Occupational Mycoses. Philadelphia, PA, Lea & Febiger, 1983, pp 13-28.

V. H. Garrison, E. A. Shin, W. T. Foreman, D. W. Griffin, C. W. Holmes, C. A. Kellogg, M. S. Majewski, L. L. Richardson, K. B. Ritchie & G. W. Smith (2003) African and Asian Dust: From Desert Soils to Coral Reefs, *BioScience* . May 2003 / Vol. 53 No. 5, pp. 469-480.

Mr. George Allen

To: Fred Butterfield, Designated Federal Officer
EPA SAB, Clean Air Scientific Advisory Committee (CASAC)
Ambient Air Monitoring and Methods Subcommittee

From: George Allen, AAMM subcommittee member, August 7, 2004

The following are revised written comments on PM-coarse measurement methods per Dr. Hopke's request at the AAMM subcommittee meeting on PM-coarse on July 22, 2004. As requested, a copy of these comments is being sent to Dr. Phil Hopke, CASAC AAMM Subcommittee Chair. These revised comments reflect my original comments and the discussion at the July 22 AAMM meeting.

First, I wish to express my appreciation for a well designed and executed field evaluation study of a wide range of methods for coarse mode particulate mass (PM-c) and related measurements. The results from this study provide a solid base for this subcommittee's work. I encourage EPA to continue to perform meaningful monitoring method evaluation studies in the future, and to request or allocate additional funds as necessary to continue this PM-coarse work in an effort to resolve several outstanding issues. More detail on the need for further work is included below.

It must be noted up front that the results from this study are a "best case" metric because of the high level of staff skill and attention (resources) available for all aspects of this study. The same methods deployed in a large scale routine network would most likely not perform as well; this goes across the board for all methods, but perhaps more so for those that are relatively complex. It would be very informative to deploy some of these methods on a limited scale to selected SLT agencies to assess "real-world" performance.

For multiple reasons, PM-c is inherently more difficult to measure with the precision and accuracy possible for PM_{2.5} or PM₁₀ measurements. Particles substantially larger than 1 μm are more likely to be lost or to bounce on sampler inlets and surfaces; there is no truly "direct" way (simple, single measurement) to measure this size range; and PM-c concentrations are often substantially lower than PM_{2.5} in much of the U.S. Therefore, the identification of a robust "benchmark" reference method (for use in evaluation of candidate reference or equivalent methods) is critical. For this, I agree with the approach taken by Vanderpool et al. in the EPA PM-c field study where other methods are compared to a very carefully operated collocated pair of low-volume FRM samplers for PM_{2.5} and PM₁₀, with PM-c calculated by difference. With care, this approach can provide results with coefficients of variation of better than 1.5% for single sampler precision (Allen et al. 1999, JAWMA **49:PM**, 133-141). This is the least ambiguous PM-c method as well as being technically compatible with the existing PM_{2.5} FRM. There are also minimal concerns related to mechanical loss of particles from the filters (coarse mode particles are presumably better "bound" to the PM₁₀ filter substrate by fine mode particles). As such, I can not recommend any other approach for this purpose. The EPA study's experience with coarse particle loss from the R&P dichotomous sampler is a good example why

the difference method must be the “benchmark” for evaluation of other candidate reference or equivalent PM-c methods.

Specifically, the reference method should not be a manual (or sequential) dichotomous sampler. There are too many performance uncertainties in the virtual impactor approach (regardless of design), performance can vary strongly as a function of coarse to fine mode ratios and aerosol loading, particles on the coarse filter can be lost from the filter media after collection, and the method is inherently more complex. This does not necessarily mean that the low-vol FRM difference method is the best choice for routine deployment in large scale monitoring networks; the performance of the difference method in the EPA PM-c study was optimal because of the available resources, and would likely be significantly degraded in routine use. There is nothing inherently wrong with the reference method not being routinely used for SLT measurement programs (ozone and sulfur dioxide are existing examples).

For dichotomous samplers the EPA report notes that “Effective shipping protocols resulted in negligible particle loss during transport...”. It is not clear to what extent particle loss from dichotomous sampler coarse channel filters is still a concern under less than ideal shipping conditions; e.g., should special shipping protocols be required for dichotomous filters to prevent particle loss? This has been shown to be a problem in the past, in the early 1980's when a 15 μm inlet was used (Dzubay JAPCA **33**:7, Spengler JAPCA **33**:12). Despite the EPA report’s claim of negligible particle shipping loss, the work to-date by EPA does not rule out the potential for coarse dichotomous filter shipping loss. The EPA report concludes that the rough filter handling action of the R&P sequential filter dichotomous sampler was mechanically “knocking off” up to 20% of the coarse filter mass, presumably leaving little mass to be further lost during shipping. The winter Phoenix study is too limited to be used to demonstrate resolution of dichotomous sampler coarse particle loss, both because of the small (15) sample size and the relatively small coarse to fine mode mass ratios for that study. R&P now has a revised version of the sequential dichotomous sampler that addresses this filter transport mechanism issue. The best way to test both the sampler design changes and potential for shipping loss is to return to a site like Phoenix in the summer with both high coarse mass concentrations AND a high coarse to fine mass ratio.

A parameter of concern for PM-c integrated filter samplers is the ability of the sampler to routinely produce field blanks with minimal mass gain. Existing EPA PM_{2.5} FRM regulations allow up to 30 μg mass gain on field blanks. For areas with PM-c means in the range of 10 to 15 $\mu\text{g}/\text{m}^3$ (much of the eastern US), a PM-c sampler (difference or dichotomous approach) that was close to the limit of meeting this specification would produce data with degraded precision (a high blank value implies a variable blank value). It is recommended that the field blank limit of 30 μg be reduced to half that value or less. With proper sampler design and filter handling procedures, mean field blank values of less than 5 μg are readily achievable.

The issue of using a design-specification reference method approach (as was done for the PM_{2.5} FRM) or a performance specification approach (used for the PM₁₀ HiVol, resulting in some method biases) for approval of candidate PM-c methods was discussed. I support a performance specification approach, since it widens the scope of technologies that could be candidates for approval. This approach does require careful consideration of the required tests to assure that a method works well over a wide range of situations, including differences in sample

chemical composition, wind speed, ambient temperature, mass loading, fine to coarse mass ratio, etc.

The ability of a method to meet multiple monitoring objectives (beyond only NAAQS compliance) is important. However, no single PM-c method will meet all the goals noted in Rich Scheffe's June 18th 2004 memo; that would take a mixture of integrated filter and real-time (continuous or semi-continuous) methods as is presently implemented in the PM_{2.5} and STN networks. The integrated (filter-based) methods (FRM-difference or dichotomous samplers) are appropriate for trends, compliance with NAAQS, and chemical speciation. Other than the differences noted above, the only additional comment in this context is that PM-c speciation is more readily done with a dichotomous sampler, since the coarse-channel filter is mostly coarse-mode aerosol; this improves the performance of XRF and some other analytical methods, and also allows more precise measurement of PM-c chemical components that are primarily in the fine mode (a difficult task when doing PM-c speciation by difference). It would also be useful to have more information on the virtual impactor design used in the R&P 2500D dichotomous sampler; it is an "EPA-design" according to the sampler's manual, but the pictures in the 2500D manual appear to be the classic Loo and Cork design used in earlier commercial dichotomous samplers, which is not an "EPA-design".

The "continuous" (real-time) methods are essential for public reporting (AQI, AIRNow, media alerts, etc.), health-effect assessment of PM-c on a sub-daily time-scale, identification of short-term events, and assessment of diurnal patterns (a useful tool in source identification). These methods also have the potential to provide much more detailed data at a lower overall operational cost. I support EPA's goal of wide deployment of continuous methods for PM, but only as the available technology permits collection of data of useful quality. One of the few limitations of the available material from the EPA PM-c field study is the lack of information on sub-daily precision for any of the continuous methods. It would be expected that there could be a substantial range of performance (precision) on an hourly time-scale across the three continuous methods, but this can not be evaluated at present. I would like to see a summary of precision within and across the continuous methods for one and four hour means.

The following comments are on the study's continuous methods and are therefore limited to the context of daily mean metrics, even though that is only part of the stated monitoring objectives for the PM-c program. None of the tested continuous methods were without significant performance limitations, although there are promising candidates.

R&P CM TEOM®. For all sites except Phoenix/Summer, this method read substantially lower (20 to 30%) than the reference method, but was well correlated with the FRM difference method across all sites; this might in part be explained by a 9 µm inlet cut point on the non-standard 50 LPM PM-10 inlet. However for the Phoenix/Summer test period, this method read 5% higher than the reference (a large intercept but a slope consistent with the other test sites); no rationale is given for this in the EPA report. The Phoenix/Summer test period was notable for its very high coarse mass concentrations and coarse to fine mode PM ratios, which may be associated with the different PM-c TEOM results there. However, it is worth noting that if part of the explanation for why the R&P CM TEOM method read lower than the reference at the other three test locations is the low inlet cut point size, that is potentially inconsistent with the Phoenix/Summer results. Analysis of the APS coarse mode size distribution data may help

resolve this question. Other aspects of the PM-c TEOM method are desirable, but for it to be given serious consideration as an acceptable continuous PM-c method the reason for the difference in response relative to the reference method across the four study sites would need to be identified and corrected. It is not the low bias at three of the test sites that is of concern here. It is sufficient for a method to be consistent in its response relative to the reference method across sites and seasons, although it also helps promote confidence in a method to understand why any differences exist. As a first step in this process, I would suggest a thorough and independent laboratory evaluation of the performance of the virtual impactor used in the R&P CM TEOM - particle loss, collection efficiency, and calculated enrichment factor - for a wide range of aerosol sizes (0.2 to 15 μm). Any PM-c method that does not use the classic "FRM" PM10 low-vol inlet needs to have the inlets' performance adequately tested, including tests of the inlet's aspiration/penetration curve at various wind speeds. The R&P CM TEOM 50 LPM inlet has recently been redesigned to better match the size cut of the FRM 16.7 LPM PM10 inlet; ideally both the new and old design should be characterized. If a rigorous wind-tunnel test is not possible, one simple way this could be done would be to run the FRM PM10 inlet and both (old and new) R&P coarse mass TEOM PM10 inlets as lowvol PM10 samplers, in a windy and high coarse to fine mass mode ratio environment such as Phoenix in the summer. Finally, it would be informative to revisit Phoenix during the summer with the redesigned PM-c TEOM PM10 inlet to determine if the method still has a different response relative to the FRM difference method at this site as seen in the first Phoenix summer tests.

Tisch/Kimoto SPM 613D beta dichot. This method performed reasonably well across all test periods for PM-c but was unacceptably variable for PM2.5 or PM10 for various reasons, possibly a combination of poor virtual impactor design (excessive coarse particles in the fine mode channel) and inadequate control for particle-bound water. This defeats part of the purpose of a dichotomous sampler (to provide data for both fine and coarse mode PM). The Tisch beta dichotomous method is also likely to have the worst sub-daily PM-c precision of the 3 continuous methods tested, although those data are not currently available. As with the R&P CM TEOM, if the Tisch continuous dichotomous sampler is to be considered as a viable method, the virtual impactor performance needs to be thoroughly characterized to help understand why the method performs as it does.

TSI model 3321 APS. The TSI time of flight method seemed to work best in Phoenix, regardless of season. Its PM-c correlation there was excellent, although there was a large bias (approaching a factor of 2). The PM-c correlation with the reference method at the other two sites was marginal at best. The reason for the large bias remains unexplained even after substantial efforts by the manufacturer, although this bias may be consistent with other field and lab work; it may be due to losses in the sampling system or as EPA presented at the AAMM meeting to the "shape factor" of typical coarse mode aerosols (the APS measures aerodynamic diameter, but assumes spherical shape when converting the number data to estimated mass). The ability of the APS to provide detailed size distribution along with a PM-c measurement is a desirable feature.

To summarize, I suggest additional field testing be done at two sites: revisiting Phoenix in the summer, and an eastern US site also during the summer. These two scenarios provide different but challenging environments for PM-c methods. A Phoenix summer test (NOT winter) would show how the R&P TEOM coarse mass method works with a redesigned PM10 inlet -- does it still perform differently in the Phoenix summer environment compared to the other sites? This

site in the summer could also properly address the coarse particle mass loss issues with the revised R&P dichotomous sampler (both filter handling and shipping loss). A summer eastern US location with fine to coarse mode mass ratios of 2 or more (common to much of the eastern US) would provide a new challenge to all the coarse mass methods - both integrated and continuous, and provide more information on the performance of the redesigned R&P TEOM continuous coarse monitor (new PM10 inlet, possibly run at 30°C, etc.). None of the EPA study test sites to date had fine to coarse ratios greater than about 1; as this ratio increases there is the potential for PM-c data quality degradation for ALL of the tested methods except the APS. Any interference from the 4% of PM-fine present in the PM-c TEOM method would be most noticeable at a site with high fine to coarse PM ratios and high dewpoints. For all of these additional tests, it is critical that they be performed during the summer, not the winter; the desired fine and coarse PM characteristics and relevant meteorology are a strong function of season.

There was some discussion at the AAMM meeting of the need to determine the spatial characteristics of PM-c in advance of network design and deployment, since PM-c has the potential to be much more spatially variable than either PM2.5 or PM10. The concept of performing some saturation studies to assess this issue was proposed. If this is done it is important to use methods that produce PM-c data with good precision; some of the commonly used PM saturation study methods have insufficient precision for this use. There are also some existing studies that did PM-c by difference with several sites across urban areas; one such multi-city study was funded by EPA and done by HSPH in the mid-1990's (Suh et al. 1997, *Environ. Health Perspect.*, **105**:826-834). Data from these existing studies may be useful in this context.

The EPA DQO tool is an interesting and potentially useful approach in assessing the impact of various changes to a sampling program on its ability to identify uncertainties in compliance with a "bright line" NAAQS value. This is desirable from a compliance perspective, but does not address data quality from another critical need: use of PM-c data in multi-variate health effect studies concurrently with PM2.5 or other pollutants such as CO or NO2 or O3. There is a need for PM-c data with sufficient precision so that acute (time-series) health-effect models can properly use PM-C without precision-related biases. Because of the more complex nature of PM-c measurements as noted above, most existing PM-c data have relatively poor precision compared to other NAAQS pollutants. This can bias the health-effect estimate towards the more precisely measured pollutant (White, *JAWMA* **48**:454-458). Most NAAQS pollutants have precision for daily metrics that are 2 to 5 times or more better than historical PM-c precision. If we are to make progress in properly assessing the acute health effects of PM-c, it will be essential to generate data in compliance networks with sufficient precision to minimize this source of model bias. It remains to be seen if current technologies can achieve this goal in the context of routine use in SLT networks. As noted in my opening comments, to the extent that these simulations rely on the results from this field study, the results may not reflect the performance of these methods when used in routine monitoring networks run by SLT agencies. This may limit the ability of a DQO tool to properly assess expected network performance when using the results of this study as input. Finally, the use of PM-c data generated by difference from HiVol PM10 and LowVol PM2.5 to estimate data quality has limited value, since these two methods are sufficiently different to introduce various artifacts in the difference PM-c data.

Dr. Judith Chow

July 19, 2004

To: Fred Butterfield, Designated Federal Officer, Clean Air Scientific Advisory Committee (CASAC)

From: Judith C. Chow

Cc: Phil Hopke, CASAC Ambient Air Monitoring and Methods (AAMM) Subcommittee Chair

Subject: CASAC AAMM Subcommittee Consultation on PM_{coarse} Method Evaluation

Subcommittee members were asked to comment on the strengths and weaknesses of candidate methods tested (Question 1), relate these test samplers to multiple monitoring objectives (Question 2), and evaluate the process used to develop PM_{coarse} data quality objectives (DQO) (Question 3). Substantial efforts have been made in the multi-site evaluation and in the development of PM_{coarse} DQO. These activities provide a good starting point; however, additional documentation is needed and more testing may be required to fully address these questions. The following observations and suggestions address issues regarding: 1) study design; 2) selection of samplers; 3) sample processing and validation; 4) data analysis; 5) recommendations on further method characterization and development by manufacturer; and 6) development of PM_{coarse} DQO.

1. Study Design

No single measurement method can meet all the monitoring objectives for compliance, diurnal variations, public alert, chemical characterization, source attribution, and long-term trends. A combination of both integrated and in-situ continuous samplers is needed. The major challenges in PM_{coarse} measurements are: 1) sensitivity of sampling efficiency to inlets with different sampling effectiveness between 2.5 and 10 μm and 2) large spatial variations. The major features of mass distributions of particle sizes found in the atmosphere consist of multiple modes. The peak of coarse modes may shift between ~6–25 μm (Lundgren and Burton, 1995). As pointed out in Chow (1995), small shifts in the 50% cut-point of PM₁₀ samplers will have a large influence on the mass collected because coarse mode often peaks near 10 μm . On the other hand, a similar shift in cut-point near 2.5 μm will have a smaller effect on the mass collected, owing to the low quantities of particles in the 1–3 μm size range. See Watson et al. (1983) and Wedding and Curney (1983) for more on this topic.

In the report (Vanderpool et al., 2004), a table that documents inlet type (e.g., operating principle), 50% cut-point, slope, and a reference to the tests is needed to provide readers with an overview of the inlet's cut-point and the sharpness of the curve. (See attached example in Table 1 by Watson and Chow, 2001.) In addition, a diagram of the sampler layout on a horizontal scale (assuming the same layout is used for every location tested) is needed to evaluate collocated precision (e.g., two samplers next to each other might agree better than a pair at opposite ends of the platform). Eight PM_{coarse} samplers were installed inside the motor home with extended downtubes. What is the inlet height above ground level? Will the large sample air travel distance,

residence times, and particle deposition between the inlet and the filter/sensor result in variations in PM concentrations? On-site meteorological conditions are useful for the interpretation of the data. It is not clear from the report where the location of the meteorological sensors is with regard to the testing platform and what the averaging times are for wind speed (WS) and wind direction (WD). Gustly winds during the test period are of importance for explaining abrupt concentrations in the continuous PM measurements.

Given the difficulties in selecting proper locations for testing, some background information regarding emissions, seasonal variations, and PM composition in each sampled area would be helpful. PM_{coarse} may be higher during fall, as agricultural activities increase. Additional tests under high-wind periods and during fall are recommended.

2. Sampler Selection

Five types of samplers (two types of filter samplers and three continuous methods) were tested, which included: 1) BGI, Rupprecht & Patashnick (R&P), and Thermo-Andersen Federal Reference Methods (FRMs); 2) R&P 2025 dichotomous sampler (dichot); 3) R&P PM_{coarse} tapered element oscillating microbalance (TEOM); 4) Tisch Dichot beta attenuation monitor (BAM); and 5) TSI aerodynamic particle sizer (APS). These included two types of PM_{10} inlets, both of which are based on particle impaction for 10 μm size cuts at 16.67 and 50 lpm. It appears that the FRMs, R&P 2025, Tisch, and APS used a so-called standard PM_{10} inlet (assumed to be the louvered SA246 inlet) at 16.67 lpm. Only the R&P PM_{coarse} TEOM used a 50 lpm inlet. Virtual impactors are used in three samplers; the major and minor flows varied from 15 and 1.7 lpm in the R&P 2025, 15.2 and 1.5 lpm in the Tisch, and 48 and 2 lpm in R&P TEOM, respectively. Besides the standard PM_{10} and WINS inlets, are sampling effectiveness curves available for these other inlets? Will particles overload in a higher-flow-rate sampler in a polluted environment over the 30-day testing period? Either inlet overloads or high wind speeds may result in changes in sampling effectiveness curves. In addition, size selective properties of inlets are not necessarily the same for all sampled particle sizes and for all conditions of the inlet.

The selection criteria listed in Vanderpool et al. (2004) are reasonable, except that the $PM_{2.5}$ FRM standard cassette does not seem to be a necessary requirement for a candidate PM_{coarse} sampler. The IMPROVE sampler, for example, that has performed $PM_{2.5}$ and PM_{10} monitoring since 1988 does not use a standard FRM cassette, yet its measurements will be used to establish a five-year baseline for the Regional Haze Rule (Watson, 2002). Why wasn't the URG FRM included? For the integrated samplers, only the R&P 2025 sequential dichot sampler was tested in addition to FRMs. Why wasn't the Thermo-Andersen manual dichot (the "original" dichot) tested? This manual dichot sampler was the EPA-designated reference method (RFPS-0389-073, Federal Register, Vol. 54, p. 31247, 7/27/89). It is commercially available and was used extensively during the 1980s for the U.S. EPA's inhalable particle network. It was tested with prototype FRMs for $PM_{2.5}$ (Pitchford et al., 1997). Although there is no $PM_{2.5}$ FRM for high-volume samplers (HIVOL), they are commercially available. Also commercially available are battery/solar-powered mini-volume $PM_{2.5}$ and PM_{10} samplers (AirMetrics and BGI). High-volume samplers provide large mass loadings for organic speciation and mini-volume samplers are useful for better understanding spatial variabilities and conducting indoor/outdoor exposure assessments. Given that EPA is re-evaluating its current network, these mini-volume and

continuous samplers can be used at a so-called Level III NCore site as part of the National Ambient Air Monitoring Strategy (U.S. EPA, 2002).

For continuous monitoring, the virtual impactor in dichot has the advantage of getting both PM_{fine} and PM_{coarse} fractions concurrently through one instrument, but it also introduces uncertainties in flow splitting and contamination of PM_{coarse} with PM_{fine} . Given that there are three PM_{10} Federal Equivalent Methods (FEM), some PM_{coarse} FEMs may be considered. Using the same difference technique as is used for the PM_{10} minus $PM_{2.5}$ FRM, these BAM and TEOM continuous methods should also be tested with PM_{10} and $PM_{2.5}$ inlets to better understand them.

Many state and local agencies may already have one or the other type of continuous $PM_{2.5}$ or PM_{10} sampler, so these comparisons can give them a better perspective on how to best use their existing resources.

To avoid the perception that EPA is favoring a certain type of sampler, a summary of surveys on the existing $PM_{2.5}$ and PM_{10} instruments used in U.S. ambient monitoring networks should be given and criteria for selection of candidate samplers should be more explicitly justified.

3. Sample Processing and Validation

For the integrated filter samplers, discrepancies between $PM_{2.5}$ and PM_{10} mass determination should be revisited regarding filter cassette shipping and handling, filter equilibrium, corrections for sample volume, and field flask shipping and handling. Much more stringent requirements in sample shipping and handling are required for $PM_{2.5}$ but not for PM_{10} . “Effective shipping protocol” (page 23 of Vanderpool et al., 2004) is mentioned without documenting the procedure.

Filter equilibration

Filter equilibration conditions for temperature and relative humidity (RH) are within ± 3 °C between 15–30 °C, and within $\pm 5\%$ between 20–45% for PM_{10} , and within ± 2 °C between 20–23 °C and within $\pm 5\%$ between 30–40%. In conducting gravimetric analysis, the number of days prior to and after sampling and the conditions for filter storage are required only for $PM_{2.5}$ and not for PM_{10} .

Sample volume construction

PM_{10} sample volumes are adjusted to sea level pressure and at 25 °C for PM_{10} , whereas no adjustment is used for $PM_{2.5}$. See page v of Appendix 3A in Attachment 3, where the standard condition is used in the analysis (meaning the volume used in the concentration calculation was adjusted to 1 Atm and 25 °C).

Field blanks

It is not clear if field blanks were collected during the 30-day comparison and if blank subtraction was performed. The need for field blank collection and correction should be investigated. Passive deposition of PM_{coarse} could be significant in a dusty environment.

4. Data analysis

Vanderpool et al. (2004) detailed the test performance and test results. However, much of the information is imbedded and it is not easy to make equivalent comparisons. The following information is recommended:

- A table of inlet specifications (as suggested in Section 1); a table providing sampling effectiveness slopes and cut points with reference to the detailed test report (if it is available) and the sampling system, such as measurement principles, particle size, inlet type, inlet and sampling surface, flow rate, filter holder/type, type of flow control, and its minimum averaging time and detection limits.
- Provide time series plots for every testing period on all instruments in an Appendix to allow a cross-comparison over time and location.
- Tabulate all statistical comparisons. Linear regression statistics provide only part of the information; they do not specify outliers or distribution. Mathai et al. (1990) used several performance measures to establish the equivalence, comparability, and predictability of the collocated measurements. For example, percent distribution of uncertainty intervals is needed to better understand the differences between the samplers.

Criteria for invalidating the data need to be clarified. For instance, on page 6 of Attachment 3, the precision of the R&P dichot is 3.8% (Table 3) but mean bias ranges from -9.8% to 12.5%. Notice that total data points in the R&P dichot are 47 compared with 90 for the Tisch.

Using the USC prototype during the January 2004 test appears to be an afterthought. What's the difference between the R&P PM_{coarse} TEOM and the USC TEOM? What is the basis for concluding that the two instruments are identical based on 15 days of measurements at one location?

What is the purpose of including APS while no particle size data is given? There are several types of particle size methods available commercially (e.g., Grimm OPC, Dekati ELPI, Climet OPC, MSP wide-range size classifier). There is no advantage to adding a PM_{10} inlet, assuming the same density of 2g/cm^3 , and comparing PM_{coarse} mass with others. The conclusion that APS has an acceptable level of precision is not supported by the data presented. Mean bias in APS varied between 50-54% at all sites (see Table 2 of Attachment 3) with overall precision of 19.5% (Table 3 of Attachment 3). Note that large discrepancies (>2 fold) were found on days 10 and 12 in Figure 14 (p. 22 of Vanderpool et al., 2004) for the collocated APS in Gary, IN. There appears to be no consistent relationship between PM_{coarse} mass, and the differences between filter measurements and APS measurements seem larger with higher concentrations.

APS in this comparison is not used properly. Why not examine the particle size distributions for each location and for differed meteorological characteristics? What are the situations of particle size distributions under average and extreme conditions? These data may shed light on the discrepancies among candidate samplers.

5. Recommendation for Further Testing

- For collocated PM_{2.5} and PM₁₀ FRMs, these inlets have undergone numerous tests and appear to be well documented. The PM_{coarse} determined by difference should be called a “benchmark” sampler in order not to confuse it with the term FRM. Testing of other high-volume and mini-volume samplers should also be considered.
- The R&P 2025 sequential dichot may be compared to manual Andersen dicot that was designated as PM₁₀ FRM. Various virtual impactors need to be characterized for sampling effectiveness. Common problems in sequential sampling systems are 1) the reliability of the transfer mechanism and 2) leakage.
- Other PM₁₀ and PM_{2.5} BAMs and TEOMs should also be included to better understand how the differences in PM_{coarse} result from inlet function and from different measurement principles.

The Tisch dichot appears to show the potential for both hourly measurements and their chemical composition if: 1) the filter punch can be advanced on an hourly or 2-3 hour basis; 2) the filter media are suitable for subsequent elemental and ion analysis (note that low hygroscopicity polyfon is listed on page 5 of the report of Vanderpool et al., but a glass-fiber filter is shown in Vanderpool’s presentation); and 3) a light absorption measurement is used as a surrogate for black carbon. (This is available on some of the Kimoto units.)

Even though the three Tisch instruments show acceptable precision (10%), several improvements need to be made and tests need to be conducted:

- What are the standards used for calibration? Why not calculate volume flow from the mass flow control by measuring temperature and pressure? Beta attenuation is sensitive to RH for hygroscopic particles. Maintaining 25 °C with an external heater downstream of the inlet will not minimize the RH interference since atmospheric RH can still be >65% at 25 °C. (Consider Atlanta, GA, or Riverside, CA, during summer.) The sampler should be heated only when RH is >65%. Smart heaters can be used to achieve RH < 65%. Beta attenuation depends on the atomic number and, hence, particle composition of the aerosol (Jacklevic et al., 1981). Since the dominant atomic numbers and chemical compositions of PM_{fine} and PM_{coarse} differ, their calibration constants are probably different. To eliminate problems with the flow splitter, another option is to use well-tested PM₁₀ and PM_{2.5} inlets separately since both channels operate independently.

6. Development of PM_{coarse} DQO

The approach taken here is highly technical, but it appears to be well grounded in statistical theory. The description of the DQO tool would be improved if it more clearly stated its assumptions in plain language, probable deviations from these assumptions that are likely to occur in practice, and the magnitude of uncertainty introduced by various (but probable) deviations.

The effect of sampling frequency on uncertainties in the 98th percentile and annual average is a useful feature of this analysis. A less precise indicator (e.g., from the TEOM and BAM) might provide a more certain estimate owing to its ability to sample every day.

The application applies to measurements at a single site used to determine compliance with assumed annual and 24-hour forms of a PM NAAQS. The PM_{2.5} annual average standard uses a spatial average of measurements at community representative sites. Sites with averages that differ from the spatial average by more than $\pm 20\%$ are removed from the spatial average and examined against the 24-hour 98th percentile average, along with source-oriented monitors. This form was intended to better represent the uncertainty of spatial variability and to better represent population exposure. A larger number of less precise monitors might provide a better estimate of effects-related exposures if they could be deployed for the same cost as a certain number of integrated filter samplers. Decisions such as these (i.e., how do I obtain greater benefit for the same cost) are likely to become more common as EPA's National Ambient Air Monitoring Strategy (U.S. EPA, 2002) is implemented.

The DQO application does not examine other uses of PM data to improve public health. These should also be assessed based on a given investment in monitoring resources rather than on an individual site and sampler basis. It may be that some combination of different monitoring methods would best achieve the objectives for the same resource investment.

Data uses might include:

- Broadcasting air quality alerts and perfecting air quality forecasts (Stockwell et al., 2002). This would require real-time hourly monitors. The uncertainty of the measurement would be balanced against the cost of the network. It is probable that spatial coverage for the same amount of resources would be the deciding point. Adding an additional monitor probably provides more benefit than a 5% improvement in precision.
- Source identification. Time resolved measurements can be correlated with wind direction and identifiable events. Filter samples can be analyzed for more specific source markers. Some combination of different sampler types for the same resources might allow both.
- Source amelioration and emission inventory. A real time measurement would allow an inspector to identify a problem and immediately remediate it. Over the long-term, examination of deviations from diurnal patterns would help to identify sources that are not included in the inventory and when they are most likely to occur.

- Future health relationships. Outdoor human exposure could be better quantified with shorter duration, more frequent, and spatially diverse sampling. Presumably, epidemiological relationships between adverse health effects and monitoring results would be better with more frequent and spatially diverse measurements from continuous monitors. They would also be enhanced by more specific chemical compositions that are currently only available from integrated filter samples.

To improve the analysis, the DQO model needs to incorporate spatial distribution as well as optimization criteria for additional air quality monitoring objectives that are above and beyond the determination of compliance with NAAQS. The overall objective should be to optimize monitoring networks to improve public health at the least cost to the economy and within a given allocation of monitoring resources.

References

- Chow, J.C. Critical review: Measurement methods to determine compliance with ambient air quality standards for suspended particles; *JAWMA* **1995**, 45(5), 320-382.
- Lundgren, D.A.; Burton, R.M. Effect of particle size distribution on the cut point between fine and coarse ambient mass fractions; *Inhal. Toxicol.* **1995**, 7(1), 131-148.
- Jaklevic, J.M.; Gatti, R.C.; Goulding, F.S.; Loo, B.W. A beta-gauge method applied to aerosol samples; *Environ. Sci. Technol.* **1981**, 15(6), 680-686.
- Mathai, C.V.; Watson, J.G.; Rogers, C.F.; Chow, J.C.; Tombach, I.H.; Zwicker, J.O.; Cahill, T.A.; Feeney, P.J.; Eldred, R.A.; Pitchford, M.L.; Mueller, P.K. Intercomparison of ambient aerosol samplers used in western visibility and air quality studies; *Environ. Sci. Technol.* **1990**, 24(7), 1090-1099.
- Pitchford, M.L.; Chow, J.C.; Watson, J.G.; Moore, C.T.; Campbell, D.E.; Eldred, R.A.; Vanderpool, R.W.; Ouchida, P.; Hering, S.V.; Frank, N.H. *Prototype PM_{2.5} federal reference method field studies report – An EPA staff report*; U.S. Environmental Protection Agency: Las Vegas, NV, 1997. <http://www.epa.gov/ttn/amtic/pmfrm.html>.
- Stockwell, W.R.; Artz, R.S.; Meagher, J.F.; Petersen, R.A.; Schere, K.L.; Grell, G.A.; Peckham, S.E.; Stein, A.F.; Pierce, R.V.; O'Sullivan, J.M.; Whung, P.Y. The scientific basis of NOAA's air quality forecasting program; *EM* **2002**(Dec.), 20-27.
- U.S.EPA. *National ambient air monitoring strategy (Second draft)*; U.S. Environmental Protection Agency: Research Triangle Park, NC, 2002. <http://www.epa.gov/ttn/amtic/files/ambient/monitorstrat/compms.pdf>.
- Vanderpool, R.; Ellestad, T.; Handley, T.; Scheffe, Richard, Hunike, E.; Solomon, P.; Murdoch, R.; Natarajan, S.; and Noble, C. *Multi-Site Evaluations of Candidate Methodologies for Determining Coarse Particulate Matter (PM_c) Concentrations*; Draft Report; 2004.
- Watson, J.G. Visibility: Science and regulation; *JAWMA* **2002**, 52(6), 628-713.

Watson, J.G.; Chow, J.C. Ambient air sampling, in *Aerosol Measurement: Principles, Techniques, and Applications, Second Edition*, 2nd ed., Baron, P., Willeke, K., Eds.; John Wiley & Sons: New York, NY, 2001, pp. 821-844.

Watson, J.G.; Chow, J.C.; Shah, J.J.; Pace, T.G. The effect of sampling inlets on the PM₁₀ and PM₁₅ to TSP concentration ratios; *JAPCA* **1983**, 33(2), 114-119.

Wedding, J.B.; Carney, T.C. A quantitative technique for determining the impact of non-ideal ambient sampler inlets on the collected mass; *Atmos. Environ.* **1983**, 17, 873-882.

Table 1. Size-selective inlets and characteristics for ambient aerosol sampling (Watson and Chow, 2001).

Name, Manufacturer, References	Inlet ID: d ₅₀ (µm), Slope, Flow (l/min)	Flow Rate (L/min)	Description and Comments
Impactor			
Airmetrics Minivol Impactor (ARM) (Turner, 1998; Wiener et al., 1992)	MV10: ~10,NA,5 MV2.5: ~2.5,NA,5	5 5	Available in machined polymeric propylene plastic or machined aluminum. PM ₁₀ and PM _{2.5} inlets are used in series for PM _{2.5} sampling. Apiezon vacuum grease dissolved in hexane is pipetted onto impaction surfaces before each sample to minimize re-entrainment.
Harvard Sharp Cut Impactors (ADE) (Marple et al., 1987; Turner et al., 2000)	MST123: 1,1.22,23 MST24: 2.5,1.02,4 MST210: 2.5,1.06,10 MST220: 2.5,1.25,20 MST104: 10,1.11,4 MST1010: 10,1.09,10 MST1020: 10,1.06,20		Machined aluminum.
URG Impactors	URG25AA:1,NA,4 URG25PAB:10,NA,4 URG30DBE: 10,NA,16.7	4	Used for personal particulate sampler and indoor air monitoring.
URG-2000-30DBE Impactor	10 µm	16.7	Anodized PM ₁₀ inlet.
URG-2000-30DBF Impactor	10 µm	16.7	Used on URG dual sequential fine particle sampler and URG weekly air particle sampler.
URG-2000-30DBN Impactor	10 µm	32	Versatile Air Pollutant Sampler (VAPS) Teflon-coated PM ₁₀ inlet.
URG-2000-37F Impactor	2.5 µm	2	Used with 37 mm impactor filter pack.
URG-2000-25A Impactor	2.5 µm	4	Used for personal particulate sampler and indoor air monitoring.
Impactor/Elutriator			

Name, Manufacturer, References	Inlet ID: d ₅₀ (µm), Slope, Flow (l/min)	Flow Rate (L/min)	Description and Comments
Andersen (GRA) hivol PM ₁₀ (Kashden et al. (1986);McFarland et al., 1984); Ranade et al., (1990);Wedding et al. (1985)	G1200: 9.7,1.4, 1,133	1,133	Anodized spun aluminum with a single stage of opposing jets. The body is hinged to facility cleaning and re-greasing of the removable impaction plate that is sprayed with an aerosol adhesive after cleaning. The G1200 was preceded by the SA-320 single stage PM ₁₅ inlet and the SA321A and SA321B dual stage PM ₁₀ inlets that are no longer sold but may still be in use. It is not entirely clear which sampling effectiveness tests apply to each of these inlets.
Andersen (GRA) medvol PM ₁₀ Olin and Bohn (1983)	SA254I: 10,1.6,113	113	Spun aluminum with ten impactor jets and a central elutriation tube. The inlet can be disassembled for cleaning. The SA254I was preceded by the SA254, or “Blue Head” owing to its enamel painting that was nearly impossible to disassemble for cleaning.
Andersen (GRA, R&P, URG) lovol “Flat Top” PM ₁₀ McFarland et al. (1978); Van Osdell and Chen (1990); Wedding et al. (1980)	SA246B: 10.2,1.41, 16.7	16.67	Machined aluminum with three parallel impactor tubes and a central particle elutriator tube. Rain drops are blown into the inlet beneath the flat top and accumulate on the impaction surface. Water exits through a small drain attached to a bottle on the outside of the inlet. The top unscrews for cleaning impactor surfaces.
FRM (BGI, GRA, R&P, URG) lovol “Curved Top” PM ₁₀ Federal Register (1997)	Curved Top PM ₁₀ : 10,NA,16.7	16.67	Same materials and design as the SA246B but with a top that curves over the inlet bug screen to minimize the entry of windblown raindrops.
EPA (BGI, GRA, R&P, URG) Well Impactor Ninety Six Federal Register (1997); Kenny et al. (2000)	WINS: 2.48,1.18,16.7		Machined aluminum well with a detachable impactor jet. The impaction surface consists of a 37 mm quartz fiber filter immersed in 1 ml of vacuum pump oil to minimize particle re-entrainment over multiple sampling days between cleaning.
Virtual Impactor			
Andersen (GRA) dichotomous virtual impactor, McFarland et al. (1978)	SA241: 2.5 µm,NA	16.67	RFPS-0789-073. Designated for PM ₁₀ dichotomous sampler only.
VAPS (URG) Virtual Impactor	VAPSVI:2.5,NA,32	32 32	
Cyclone			
Wedding (GRA) IP ₁₀ (Wedding et al., 1982)	IP10: 9.6,1.37,1133	1,133	RFPS-1087-062. Inlet cleaning port on top of inlet.
Andersen SA246B2.5	2.5 µm 2.4	16.67	Typically used with SA246B PM ₁₀ inlet.
Andersen 3.68 Cyclone (modified AIHL)	2.7 µm 1.16 2.3 µm 1.18	24 28.1	Used on Andersen RAAS speciation sampler.

Name, Manufacturer, References	Inlet ID: d ₅₀ (µm), Slope, Flow (l/min)	Flow Rate (L/min)	Description and Comments
BGI GK-2.69 Cyclone	10 µm 4.0 µm	1.62 4.2	PM ₁₀ /thoracic – oil mist. High flow respirable – silica.
BGI SCC-1.062 Cyclone	1.0 µm 1.21 2.5 µm 1.20 4.0 µm 1.22	3.5 1.5 1.05	For indoor air quality PM ₁ use. The inside diameter of the cyclone is 1.062 cm. For indoor air quality PM _{2.5} use. For indoor air quality respirable use.
BGI SCC-2.229 Sharp Cut Cyclone	1.0 µm 1.17	16.7	For use with the BGI PQ200. The inside diameter of the cyclone is 2.229 cm.
BGI SCC-A and SCC-B Sharp Cut Cyclones	2.5 µm 1.19	16.7	For use with the BGI PQ200. Equivalent to EPA WINS.
IMPROVE Cyclone (modified AIHL)	2.5 µm	22.7	Modified Air Industrial Hygiene Laboratory cyclone.
Met One SCC-1.118 Sharp Cut Cyclone	2.5 µm 0.81	2	The inside diameter of the cyclone is 1.118 cm.
Rupprecht & Patashnick SCC-1.829 Sharp Cut Cyclone	2.5 µm 1.23	5	The inside diameter of the cyclone is 1.829 cm.
Met One SCC-2.141 Sharp Cut Cyclone	2.5 µm 1.24	6.8	The inside diameter of the cyclone is 2.141 cm.
MSA	0.78 µm	2	Generally used in personal sampling applications.
Sensidyne BDX99R	4 µm 1.56	1.7	Formerly the Gilian version BDX 99R. Generally used in personal sampling applications.
SKC Cat. No. 225-01-02 Cyclone	5 µm	1.9	Generally used in personal sampling applications.
URG-2000-30EHB Cyclone	1 µm	16.7	Aluminum surface with Teflon coating.
URG-2000-30EAM Cyclone	10 µm	15	Aluminum surface with Teflon coating.
URG-2000-30ENB Cyclone	10 µm	16.7	Aluminum surface with Teflon coating.
URG-2000-30EA Cyclone	10 µm	28.3	Aluminum surface with Teflon coating.
URG-2000-30ED Cyclone	2.5 µm	3	Aluminum surface with Teflon coating.
URG-2000-30EN Cyclone	2.5 µm	10	Aluminum surface with Teflon coating.
URG-2000-30EH Cyclone	2.5 µm 1.37	16.7	Aluminum surface with Teflon coating.
URG-2000-30EC Cyclone	3.5 µm	28	Aluminum surface with Teflon coating.
Stacked Filters			
Nuclepore Filters			
BGI CIS Foam	10 µm 4 µm 2.5 µm	3.5 3.5 3.5	For indoor air quality PM ₁₀ /thoracic use. For indoor air quality respirable use. For indoor air quality PM _{2.5} use.

Mr. Bart Croes

U.S. EPA's PM Coarse Methods Evaluation and Data Quality Objectives

July 22, 2004 Consultation Meeting

CASAC AAMM Subcommittee Review Comments, Bart Croes

Overall, the multi-site PM coarse sampler intercomparison and performance-based approach to determining data quality objectives (DQOs) represent an impressive initiative by U.S. EPA to take a systematic approach towards implementation of a likely coarse particle (PM_c) National Ambient Air Quality Standard (NAAQS). U.S. EPA's equal emphasis on continuous samplers is refreshing as these are a vital component of a comprehensive air quality monitoring program. Time-resolved, real-time availability of PM data are necessary for use air quality index (AQI) forecasting and burn allocations, and can lead to a better understanding of emission sources, transport, background levels, deposition, and health effects of PM. I appreciate the opportunity to comment during this intermediate stage of the process. The documents provide a thorough description of the monitoring methods and protocols for the intercomparison, clearly explain the monitoring results, and provide a reasonable rationale for the development of and inputs to the DQO software tool. I agree with the basic approach taken by U.S. EPA, and offer comments on several aspects that need further attention. My comments address the three basic questions posed by Rich Scheffe in his June 18, 2004 memo to Fred Butterfield, as well as additional issues raised by U.S. EPA staff and others at the consultation meeting.

Responses to Questions

For Questions 1 and 2 regarding the strengths and weaknesses of each method tested for purposes of a reference method or measurement principle, and to meet multiple monitoring objectives, I completely agree with the comments made by Peter McMurry (as replicated below with minor edits).

PM Monitoring Method Strengths and Weaknesses (adapted from Peter McMurry)

PM Monitoring Method	Strengths	Weaknesses
PMc FRM = PM10 FRM minus PM2.5 FRM	<ol style="list-style-type: none"> 1. Uses established monitoring equipment 2. Filters can be analyzed for particle composition 	<ol style="list-style-type: none"> 1. 24-hour time resolution 2. Expensive, manual filter analysis 3. Not useful for real-time AQI reporting 4. Involves collection of particles on a filter, rather than direct measurements of gas-borne particles
R&P 2025 Sequential Dichotomous Sampler	<ol style="list-style-type: none"> 1. Single sampler for both fine and coarse mass/composition 2. Less influenced by fine/coarse missing than PM10-PM2.5 3. Filters can be analyzed for particle composition 	<ol style="list-style-type: none"> 1. 24-hour time resolution 2. Expensive, manual filter analysis 3. Not useful for real-time AQI reporting 4. Involves collection of particles on a filter, rather than direct measurements of gas-borne particles
R&P Continuous Coarse TEOM Monitor	<ol style="list-style-type: none"> 1. Fast time response for better information on temporal exposures 2. Established track record 3. Data can be used for real-time public reporting 4. May currently have greater potential for accuracy than other methods. 	<ol style="list-style-type: none"> 1. Volatilization losses for high temperature collection 2. Apparent sampling losses for coarse particles 3. Involves collection of particles on a filter, rather than direct measurements of gas-borne particles
Tisch Inc. Model SPM-613D Dichotomous Beta Gauge	<ol style="list-style-type: none"> 1. Fast time response for better information on temporal exposures 2. Established track record 3. Data can be used for real-time public reporting 	<ol style="list-style-type: none"> 1. PM2.5 mass in poor agreement with those from other samplers 2. Involves collection of particles on a filter, rather than direct measurements of gas-borne particles
TSI Inc. Model 3321 Aerodynamic Particle Sizer (APS)	<ol style="list-style-type: none"> 1. Provides valuable supplemental information on size distributions 2. In-situ measurements of gas-borne particles (the type we breathe!) 	<ol style="list-style-type: none"> 1. Not a mass measurement method; should not be considered as such

For Question 3 regarding the appropriateness of the uncertainty estimates and completeness of the factors considered for the DQOs, my expertise is limited. Since this is a new, relatively untested software tool, the results should be reviewed by monitoring experts at U.S. EPA and State, local, and tribal (SLT) agencies to ensure that they match common sense. In addition, U.S. EPA and other staff associated with AirNow and the air quality epidemiology community should be consulted on DQOs for their uses of PM monitoring data.

Other Issues

Further analyze the sampler intercomparison data.

Suggestions for further analysis include consideration of speciation data (including anions), particle size data (from the APS), and meteorological data to further understand sampler performance and differences.

Define the potential scope of a national PMc monitoring network.

While U.S. EPA has not yet promulgated a coarse particle NAAQS, it has released a Staff Paper with a proposed range of possible standards for PM_{2.5} and PMc. As a first-order estimate, data from the existing PM₁₀ monitoring network should be compared to the proposed lower and upper ranges of the coarse particle recommendations to determine if the potential scope of a PMc monitoring network would be national in scale or restricted to a few states. In these likely nonattainment areas, PM₁₀ would primarily consist of the coarse fraction. Sites that have collocated PM_{2.5} and PM₁₀ monitors, or SLT agencies that have operated dichot samplers (see Motallebi, et al., 2003ab for California) provide more relevant data. A list and map of sites with PM₁₀ only, PM_{2.5} only, and both would be a useful summary.

Allow PM₁₀ monitors to be used to determine attainment.

U.S. EPA and SLT agencies have already invested huge resources into the current PM₁₀ and PM_{2.5} monitoring networks. Several states (i.e., California) have State ambient air quality standards for PM₁₀ and do not plan to follow U.S. EPA in adopting a coarse particle standard. Surely if a site meets the PMc standard with PM₁₀ monitoring data (uncorrected), then there is no need to deploy a PMc-specific monitor at the site.

Analyze special studies to determine spatial distributions of PMc.

A key issue for potential PMc nonattainment areas is the number of monitors that need to be sited to properly represent population exposure. The California Air Resources Board and perhaps other SLT agencies have conducted special studies. One example is a PM saturation study conducted with mini-vols during 2000 in Corcoran, an agricultural community in the San Joaquin Valley with high dust levels. Similar studies may have been conducted in Las Vegas and Phoenix.

Define the difference method as the Federal Reference Method (FRM).

It is unclear from the documentation provided to the committee whether or not U.S. EPA will allow the difference between PM₁₀ and PM_{2.5} FRMs to be defined as the PMc FRM. My recommendation is to allow this in consideration of the huge resources that have already been invested into the current PM₁₀ and PM_{2.5} monitoring networks, and the excellent precision results obtained in the multi-site sampler intercomparison study. After all, a difference method is already used to determine NO₂ levels.

Devote resources to developing a traceable standard for PM.

Problems with the TEOM (page 18, section 5.3, unit three) and APS (page 22, section 5.5, unit two) were only discovered during the intercomparison study because multiple units were carefully operated by U.S. EPA and monitoring industry experts. If the units were operating by themselves in an SLT agency monitoring station, it is unlikely that instrument drift and other

problems would have been noticed. Without the ability to challenge a PM analyzer with a known concentration of PM, all you have to verify proper operation of an analyzer is the "due diligence" of the site technician. If U.S. EPA took a dozen of the candidate samplers and sent them to 12 randomly selected SLT agencies, after four months they would get a dozen different regressions and correlations, no matter how consistent the analyzers performed in the controlled, three-city study.

Other continuous, criteria pollutant monitors (O_3 , NO_2 , CO , SO_2) are challenged with a known concentration each day (the in-station zero and span checks) and at six- and twelve-month intervals (independent transfer standards). For filter samplers the micro-balance used to weigh the filter is similarly "zeroed and spanned" with NIST-traceable standard weights each weighing session. Ozone does not come in a bottle, but accurate and precise quantities are generated on demand to challenge ozone analyzers. Resources should be devoted to research an accurate and precise PM generation system. I realize this would be very difficult to do with PM, but perhaps something similar to an aerosol inhaler (used for administering asthma medication) could be developed.

Considerations for a follow-on sampler intercomparison study.

Since U.S. EPA resources are too limited to two sites at most for a follow-on study, consider the use of a PM Supersite (i.e., Fresno with the high PM_{10} and $PM_{2.5}$ nitrate during the fall harvest season) and revisit Phoenix in the summer. SLT agencies should be involved to duplicate "real-world" operation. Perhaps the existing dichot monitor and high-volume PM_{10} and $PM_{2.5}$ monitors (that are already deployed by some SLT agencies) should be included to determine their suitability to determine if an area meets the NAAQS (e.g., they do not have negative biases).

References

Motallebi, N., C. A. Taylor, Jr., K. Turkiewicz, and B. E. Croes (2003a) Particulate matter in California: Part 1 – Intercomparison of several $PM_{2.5}$, PM_{10-25} , and PM_{10} monitoring networks. *J. Air Waste Manage. Assoc.*, **53**: 1509-1516.

Motallebi, N., C. A. Taylor, Jr., and B. E. Croes (2003b) Particulate matter in California: Part 2 – Spatial, temporal, and compositional patterns of $PM_{2.5}$, PM_{10-25} , and PM_{10} . *J. Air Waste Manage. Assoc.*, **53**: 1517-1530.

Dr. Kenneth L. Demerjian

CASAC AAMM Subcommittee Consultation on PM Coarse Methods Evaluation
July 22, 2004 Meeting at Research Triangle Park, NC
Review and Comments: Kenneth L. Demerjian

Overall the project reports (Multi-Site Evaluation of Candidate Methodologies for Determining Coarse Particulate Matter (PM_c) Concentrations; Use of a Performance Based Approach to Determine Data Quality Needs for the PM-Coarse (PM_c) Standard; General characterization of PM_c as found in the U.S., based upon data from current network of PM₁₀ and PM_{2.5} monitors) provided by OAQPS to the AAMM subcommittee were well written and very informative. This work provides an excellent foundation with regard to the measurement challenges and the issues that must be addressed to deploy a credible PM_c measurement network. OAQPS is to be commended for this initial effort to characterize several commercial samplers configured to measure PM_c and provided some basic performance statistics on their mass measurement capabilities relative to what some might argue is an arbitrary measurement standard (i.e., PM₁₀ FRM – PM_{2.5} FRM). These studies should provide the setting to assess overall performance of the sampling systems and the rationale for their observed differences. Unfortunately the current analyses fall short of the latter step, but may benefit from further analyses as the PM chemical compositional data collected as part of this program become available.

The AAMM Subcommittee was asked to focus their consultation around three major questions:

1. What are the strengths and weaknesses of each method tested in the ORD study for purposes of using it as a reference method, a measurement principle, and as a method that would provide the basis for approval of candidate reference and equivalent methods?
2. What are the strengths and weaknesses of each method tested to meet multiple monitoring objectives such as comparison to potential PM_c standards, public reporting, trends, chemical speciation, and characterization of short-term episodes and diurnal variation?
3. For the PM_c DQOs, is the process the Agency took to develop the estimates of uncertainty appropriate? Are there factors the Agency has included that should not be considered or are there other inputs that should be included?

A summary of comments & recommendations as they related to questions 1 and 2 are report in Table 1. Overall, the currently multi-site evaluations of candidate methodologies for determining coarse PM concentrations suggest that only the direct FRM PM₁₀-PM_{2.5} differencing can provide unequivocal high precision measurements of the PM coarse fraction (operational standard). Although this approach will likely have to be the operational reference method, every effort should be made to establish equivalency with one or more dichotomous filter based and continuous PM_c mass monitoring systems. Performance issues associated with these alternative methods relative to the FRM need to be addressed through more systematic studies than those presented herein. Although a reasonable start, the number of environments sampled and the sampling periods considered in this study is insufficient to draw final conclusions regarding instrumentation recommendations for a PM_c monitoring network.

TABLE I. Summary of PM Course Methods Evaluation Results

Measurement Method	PM Metric	Sampler Manufacturer(s)	Performance Metric	Comments/Recommendations
Integrated FRM	PM _{2.5}	BGI, R&P, AND	1.5% –3.4% (4)P	<ol style="list-style-type: none"> The data suggest that these methods fulfill EPA’s requirements (precedent set by PM_{2.5}) to be designated as a reference method and can be used to provide a basis for approval of candidate reference and equivalent methods. The development of absolute calibration standards for the performance testing of FRMs, remain an elusive scientific challenge to the aerosol community. Negatives: a) gravimetric methods are costly and labor intensive requiring several operator interventions; b) differencing independent (PM₁₀ and PM_{2.5}) measurements doubles operational costs and requires high precision in the independent measurements.
	PM _c		3.6% –4.1% (4)P	
	PM ₁₀		2.4% –4.7% (4)P	
Integrated Dichot, Sequential	PM _{2.5}	R&P 2025	1.3%-3.8% CV 1.00-1.09 mD/FRM	<ol style="list-style-type: none"> More data needs to be collected and analyzed to determine if the differences reported are only due to the mechanical performance of the dichotomous samplers or are also affected by the chemical composition of the ambient aerosols. Not ready as a reference method, but most likely to establish equivalency once further testing and analysis identifies source(s) of sampler bias. Negatives: a) gravimetric methods are costly and labor intensive requiring several operator interventions; Positive: a) direct collection of PM_c fraction eliminates differencing independent (PM₁₀ and PM_{2.5}) measurements reducing cost and precision demands; b) well suited for chemical speciation post analysis;
	PM _c		1.7%-4.2% CV 0.79-0.96 mD/FRM	
	PM ₁₀		1.2%-3.0% CV 0.84-0.97 mD/FRM	
TEOM	PM _c	R&P	1.7%-6.6% CV 0.69-1.05 TEOM/FRM	<ol style="list-style-type: none"> More data needs to be collected and analyzed to determine if the differences reported are only due to the mechanical performance of the sampler inlets (i.e. internal cutpoints 9 μm vs. 10μm) or are also affected by the chemical composition of the ambient aerosols. Not ready as a reference method, but most likely to establish equivalency once further testing and analysis identifies source(s) of sampler bias. Positives: a) Continuous direct mass measurement of PM_c has major cost and data utility advantages over filter based gravimetric methods (i.e. high time resolved data, real-time data display).

TABLE I. Summary of PM Course Methods Evaluation Results (continued)

Beta Attenuation	PM _{2.5}	Tisch SPM-613D	CVs not reported 1.26-1.70 Tisch/FRM	<ol style="list-style-type: none"> As with R&P PM_c TEOM more data needs to be collected and analyzed to determine if the differences reported are only due to the mechanical performance of the dichotomous sampler (Why does the R&P 2025 dichot gives different results than the Tisch Dichot?) or are also affected by the chemical composition of the ambient aerosols. Not ready as a reference method. Positives: a) Continuous direct mass measurement of PM_c has major cost and data utility advantages over filter based gravimetric methods (i.e. high time resolved data, real-time data display).
	PM _c		0.91-1.08 Tisch/FRM	
	PM		1.09-1.29 Tisch/FRM	
Time of Flight (APS)	PM _c	TSI	CVs not reported 0.42-0.62 APS/FRM	<ol style="list-style-type: none"> Many questions remain to be addressed if the APS is to be considered a viable candidate for PM_c measurements, including sensitivity to variation in PM chemical composition, relative humidity effects and variations in particle density. Not ready as a reference method. Positives: a) Continuous direct mass measurement of PM_c has major cost and data utility advantages over filter based gravimetric methods (i.e. high time resolved data, real-time data display).

The availability of chemical composition data for PM_c will prove valuable in addressing performance difference between the FRM and other methods. For example, the R&P coarse TEOM vs. FRM differences may well be explained by volatile losses of nitrate and semi-volatile organic aerosols, rather than by an inlet cutpoint issue.

The DQO process as outline in attachment 3 [“Use of a performance based approach to determine data quality needs for the PM-coarse standard”] to develop qualitative and quantitative statements regarding PM_c data, provide estimates of uncertainty and potential levels of decision error seems reasonable and should prove to be a useful tool for regulators/decision makers. It is difficult to assess, based on the write-up provided how user friendly the DQO tool is and if it will gain mainstream acceptance by decision makers. It does raise a fundamental question as to whether or not such a detailed statistical assessment creates false expectations with regard to the FRMs ability to measure the true absolute mass concentration of ambient particulate matter.

Dr. Delbert Eatough

Review Comments

Delbert J. Eatough
Professor of Chemistry
Brigham Young University

I. Multi-Site Evaluation of Candidate Methodologies for Determining Coarse Particulate Matter (PM_c) Concentrations

A. General:

EPA has conducted a multisite study to evaluate several methods for determining PM_c in anticipation of the setting of a new standard by EPA. As a result of a court decision, EPA can not set the previously anticipated PM_{2.5} and PM₁₀ standards because the fine particulate material is included in the PM₁₀ standard. Hence, if EPA decides to set a new non-PM_{2.5} standard, it must include material in a decidedly different size range. The draft document from EPA assumes this will be a PM_{10 to 2.5} size range and give this measurement the title PM_c in the document. It seems to me that since the standard is indeed new, and not just a continuation of the old PM₁₀ standard, that EPA should give some thought in the document as to justification of the supposed new size range. It's relationship to the old PM₁₀ standard is obvious. However, what is not obvious is that the PM_c as defined in the document is the best choice for a new standard.

The old PM₁₀ standard was a compromise between what was readily achievable in sampling at the time and what was known about lung deposition patterns. The fine particles are now separated out from the coarse in the path EPA is taking. Putting aside the argument of whether a 1 or a 2.5 cut is the better cut between the fine and coarse particle modes, the new PM_c standard is clearly focused on particles larger than the combustion particles and secondary products which dominate the fine particulate range. Is the sole use of the new PM_c standard to circumvent the court ruling and yet maintain the ability to track changes in what is happening relative to the old PM₁₀ standard, or is the new standard really intended to generate data which will further indicate the epidemiological need (or lack thereof) for control of coarse particles. If the latter is the case, EPA should give thought to the cut point selected for the new standard. It should represent our best understanding of lung deposition and possible exposure to coarse particulate matter. As I understand it, this point is not well met with a PM_c standard with a range of 2.5 to 10 μm.

In addition to justifying the standard from a physiological point of view, EPA should also justify the chosen standard from a sampling point of view. With the choice of an upper cut of 10 μm, EPA has virtually insured that various sampling techniques with different outlets will not be comparable. This puts the upper cut at the peak of the coarse particle size distribution for many environments and means that small changes in the inlet system will insure non-comparability of different methods. I suggest below that this effect is at the heart of the reasons for differences seen in some of the comparisons given in the manuscript. While the poor choice of a cut point may have been somewhat livable with the old PM₁₀ standard, where at least the influence of the

fine particulate mode was an ameliorating influence, the problem is much more severe with a $PM_{10\text{ to }2.5}$ standard. EPA needs to be careful that it is not boxing itself into such a sampling hole with the chosen size cut for the PMc standard that it cannot approve sampling techniques which may be much better than the “standard” of the difference measurement as outlined in the document now before the committee.

B. True Intercomparability of the Various Samplers Used:

A basic premise of the study is that the 2.5 and 10 μm cut of the various samplers are all identical. Without this intercomparability, the causes of any difference seen are not identifiable. However, the cut points, especially the more sensitive 10 μm cut points are not comparable. Arguments are made about losses in some of the systems. However, if the true cut points and the shape of the cut point curves are not known, meaningful comparisons between the various samplers are not possible. There are several points where EPA can improve the information in the report in this respect.

The characteristics of both the cut points of the FRM $PM_{2.5}$ and PM_{10} samplers have been studied and reported in the literature. EPA has defined carefully the nature of the inlet devices for both these samplers so that variability from manufacturer to manufacturer is minimal. However, the same care is not taken in the samplers chosen for incomparability in this study. Details related to this point need to be provided. Specifically:

1. The Dichot Sampler. Is the PM_{10} inlet in the dichot sampler identical to that used in the PM_{10} sampler? If not, what differences are known about the shape of the inlet curve? These differences will directly affect the total mass entering the sampler. What is known about the difference between the cut point of the fine - coarse splitter in the dichot and the WINS impactor of the $PM_{2.5}$ sampler. For both of these important cut-points, do the size distribution data obtained with the APS indicate that a specific bias would be expected in the various studies and is the nature of the bias expected to be different for the different sampling locations?

2. The Tisch Inc. Beta Gauge. The virtual impactor which makes the 10 - 2.5 cut is stated to be different for the Tisch samples. Details on the design and what is known about the shape of the curve in the 2.5 μm cut region should be given. Again, do the APS data predict any bias in the results due to the nature of the cut around 2.5 μm ?

3. R&P Continuous Coarse TEOM Monitor. The situation is even more complex for the continuous R&P instrument. Both the sensitive 10 μm cut point and the 2.5 μm cut point are different from the FRM samplers. What is known about the nature of the two curves and what do the APS data predict bias will be because of the nature of the shape of the two cut points? It is not specifically so stated, but I assume that the TX40 filter of the measuring TEOM is kept at 50 EC.

4. The APS Instrument. I am a bit confused about the need for a splitter after the PM_{10} inlet in this instrument. Why could not the APS data themselves identify the lower cut? More about the assumed density later.

5. Finally, a general comment on the various TEOM data. Were the instruments all run without the slope and intercept offsets suggested by the manufacturer so that we are looking at true results and not artificially altered results?

C. Comments on the Data Presentation:

1. It would be very useful if the figures related to the various studies were all comparable. For example, Figure 9 for the Phoenix study shows that the mass weighing for PM_{10} at the site and at RTP were comparable. However, it also includes the $PM_{2.5}$ data and clearly shows that difference around the 2.5 μm cut for the various instruments will have a minor effect on the results because fine particle were a minor contributor to the total for all data points. Similar data in the plots for results obtained at the other sites would be informative.

2. A Table of the various fine and coarse matter results would be helpful to the reader. Presently these values are scattered throughout the manuscript. The tables all refer to differences as a % of the measured as compared to the control. However, the importance of various mechanisms which can contribute to errors will be a function of the relative importance of total mass in the fine and coarse size ranges. A Table which provides these averages in a convenient place would be very helpful to the reader.

3. There is a general reliance on the presentation of regression slopes and R^2 values in the discussion of the sampler comparison. Some consideration of the calculated intercepts and the total measured mass might give a better picture. I would like to see X, Y graphical comparisons of all the data. Such a visual presentation often suggest bias or other effects not readily apparent in just linear regression results.

4. I am confused by Figure 10. The heading says the comparison is for dichot vs. FRM $PM_{2.5}$ data. However, the axis says it a comparison of FRM $PM_{2.5}$ and Dichot PMc results. Inclusion of all the linear regression analyses in the insert boxes further confuses the issue. As stated in the previous point, I would like to see comparison $PM_{2.5}$, PM_{10} and PMc X/Y plots for each of the studies. This may reveal details hidden in the limited regression results given in the paper now.

D. Specific Comments on the results:

Figure 11. How can you be certain that the difference observed in Figure 11 is due to mass loss during filter movement and not due also to the differences in the PM_{10} inlet cut point curves. The data given in Table 3 would suggest that such cut point differences were present in the data. For example, in Phoenix, where the data would be most sensitive to the nature of the 2.5 μm cut, the R&P dichot gives higher fine particulate material concentrations than the FRM. In Riverside, where the coarse particle mass averaged 30 $\mu g/m^3$, compared to a PMc average of 55 $\mu g/m^3$ in Phoenix, the dichot and FRM data differed by only about 4%. In Phoenix, the 2X higher PMc concentrations resulted in a 20 to 30% difference. Why was the mass loss also not present in Riverside. Is it possible that, while there may be some mass loss in the dichot at both sites, the difference in the shape of the coarse particles size distribution at the two sites and difference in the nature of the PM_{10} cut point for the different samplers contribute significantly to

the differences seen? The APS data may shed some light on this question. As a further example of where the APS data could be useful, on page 16 the over measurement of $PM_{2.5}$ by the dichot is attributed to incursion of coarse particles into the fine mode for this sampler. Are the known curves for the dichot and the nature of the APS data consistent with this assumption? What is the difference in the particle size distribution near $10 \mu m$ for the Phoenix, as compared to the Riverside data?

On page 18 of the text it is stated that R&P has additional data supporting loss of coarse particles during transport of the collected material. Can we get the details of these results so we can judge how applicable the studies conducted by R&P are to interpretation of the results of the EPA study reported here?

The discussion on page 19 emphasizes the importance of knowing the cut point characteristics of the various samplers. Here you attribute the difference in the R&P coarse sampler to a cut point problem. The known cut point characteristics of the various samplers really need to be detailed in the manuscript. The assumption is made that the differences between the coarse TEOM and the gravimetric results is all due to this cut point difference. Are the APS size distribution data consistent with 20 to 30% of the coarse mass being in the 9 to $10 \mu m$ range? Can any of the difference be due to semi-volatile material lost from the heated TEOM monitor? The chemical composition data may shed some light on this question. In this regard, the difference between the gravimetric and TEOM measured masses largely disappeared in the May to June Phoenix test. Do the APS data indicate no mass in the 9 to $10 \mu m$ range for these studies? Or, alternatively, is it possible that the difference in chemical composition results in less semi-volatile coarse particulate mass for the later Phoenix study?

Large biases, but with good regression slopes are seen for the Tisch Beta Gauge data. It would be very helpful to see plots of the data on which the statistics given in Table 6 are based. The assumption is made in the interpretation of the Tisch data that there was intrusion of coarse particles into the sampler fine mode. Do the APS data support the probability of this occurring. The amount of mass involved would suggest a significant tail of the coarse below $2.5 \mu m$ for this to be the case. Is the steepness of the $2.5 \mu m$ cut for the Tisch known to be much poorer than for all the other samplers to which it is compared?

E. Chemical Composition and Measured Mass:

Sufficient chemical composition data is being obtained in the various analyses to do a reasonable job for estimating mass from the composition. This analysis may shed light on whether adsorption artifacts, losses, etc. are affecting the results for any of the sampling systems. Have any data been obtained which would provide artifact free nitrate and OM concentrations for comparison with the measured mass? Are any results available which would shed light on the relative importance of semi-volatile species in the coarse particles sampled? These effects may be particularly important for the Gary and Riverside studies. I have a few additional suggestions in this regard in the last section of my comments.

The chemical composition data may also shed light on results obtained at a site such as Riverside. Substantial coarse particle nitrate could be present. This will be known when the

composition data are available. At high RH values (such as will be present at night) the uptake of water could give a very different response for the APS, as compared to gravimetric measurements which presumably excludes any water uptake. Are RH data available? During the night, periods of high relative humidity should be present in the Riverside samples.

Mention is made on page 13 of experiments with a USC prototype coarse particle sampler in the studies at Phoenix. Details of the sampler and results obtained should be given.

F. APS Results and Chemical Composition:

It is not surprising that the APS results were not in good agreement with the $PM_{10} - PM_{2.5}$ calculation of PMc. A constant density (without justification for the selected value) is used for the interpretation of all the APS data. It would be expected that the density of the coarse particles in Gary which is dominated by wind blown dust from coal pile would be very different from coarse particles dominated by suspended crustal material. For example, the ratio of APS to FRM results given in Table 6 vary from 0.4 to 0.6. Are the results off because the assumed density is not consistent with the measured chemical composition and are the different ratios, in part, due to differences in composition at the very different sites. While it may be expected that the composition (and hence density) of fine particles will be somewhat similar at each site as similar sources contribute to these fine particles, the same will not be true of the coarse particles. In fact, it appears that you have correctly chosen sites with rather different coarse particle compositions to test the samplers. Now you need to use the data you have to improve the interpretation of the APS data. In fact, it seems to me that it is only as you can bring the APS and other data together that you will really understand the data well enough to say you understand the results obtained. This will include both considerations of composition and particle cut point characteristics.

G. A Final Comment on Other Studies Which Should be Conducted:

It would be good to understand what role semi-volatile material may be playing in results obtained with the various samplers. Hopefully we will soon be to the point where we not only worry about a defensible FRM, but also about measuring the actual concentration and composition of particles in the atmosphere to assist in the interpretation of future health related studies. This point will become even more important as we put in place semi-continuous monitors to let us better understand diurnal variations and peak effects. The FDMS modification to the TEOM monitor appears to correctly measure semi-volatile species based on recent results reported by our group and others. It would be most informative to compare standard and FDMS TEOM measurements in the semi-continuous monitor (and APS results, perhaps even both heated and unheated) at the sites. If not at all sites, at least at Gary and, especially Riverside where effects might be expected.

II. Use of a Performance Based Approach to Determine Data Quality Needs for the PM-Coarse (PMc) Standard.

I do not have the expertise to completely critique this report. Better input on this document will come from others on the committee. But in general, I thought the approach was

informative and that the Gray Zone information was potentially most helpful to those who must make decisions.

Mr. Eric Edgerton

Comments on Review Material for 7/22/04 Consultation by AAMM Subcommittee

Eric S. Edgerton

The charge to the AAMM was to review documents titled “Multi-site evaluations of candidate methodologies for determining coarse particulate matter concentrations” and “Use of a performance based approach to determine data quality needs for the PM-coarse standard” and to respond specifically to three questions. Responses to the 3 questions (paraphrased) follow.

1. Strengths and weaknesses of each method tested, with respect to using it as a reference method, a measurement principle, or as a method for approval of candidate methods?

Due to the dispersion across methods, it is too early to answer this question. The only method I would discourage at this time as a “reference” is the APS, since it isn’t, and doesn’t purport to be, a mass measurement method. All other methods showed very acceptable precision, which suggests that other factors, such as inlet cutpoint(s), particle transmission and other currently uncontrolled or unknown factors are responsible for the dispersion. At this time, the only requirement I would place on a “reference” method is that it must have a well-characterized inlet.

2. Strengths and weaknesses to meet multiple monitoring objectives (e.g., comparison to NAAQS, public reporting, trends, chemical speciation)?

The continuous methods have a clear advantage over filter-based methods, when it comes to public reporting, ease of use, cost, and temporal information content. They are inferior to filter-based methods only in terms of chemical speciation.

3. For the PM_c DQOs, is the process the Agency took to develop uncertainty estimates appropriate, and are there factors that should/should not have been included?

The DQO tool is an interesting methodology for estimating overall uncertainty surrounding PM_{fine} and/or PM_{coarse} measurements, and the likelihood of Type I or Type II measurement error. It appears that the DQO tool would be used in the design of a compliance or public reporting network, but it is unclear how or whether it would be used after this (unless to verify input assumptions). Other potential applications and users should be clarified.

The Executive Summary states that “gray zones are most sensitive to population variability, sampling frequency, measurement bias and completeness.” This is true based on the input assumptions applied later in the document. It would be worthwhile to perform a sensitivity analysis to determine if this conclusion is generally applicable across the range of population parameters (e.g., as shown in Table 1).

Measurement bias of 10% may be a little optimistic for certain techniques. Inspection of Table 2, shows that only the “reference” method (FRM) consistently performed to this level. This raises two questions: first, what would the curves look like for method-specific bias estimates of say 15-20%?; and 2) would the agency consider an FEM approach to adjust for bias? Intra- and inter-method precision data suggest this might be a viable option (assuming future field tests do not show convergence of results).

Finally, it is unclear how PM_{fine} in the minor flow of the TEOM and Tisch units figures in the DQO calculation. For the Tisch unit, it seems as though the calculation should be similar to the dichot and both should be affected by the additive errors of two methods. For the TEOM, it may be argued that PM_{fine} in the minor flow is insignificant, because of the particle concentrator inlet. This is certainly true for sites with high PM_{coarse}/PM_{fine}, but perhaps not for sites with low PM_{coarse}/PM_{fine}.

Mr. Henry (Dirk) Felton

Comments for AAMM meeting to discuss PM Coarse Methods Dirk Felton July 15, 2004 Revised July 26, 2004

General Comments

States need both 24-Hr filter samples for Speciation and Continuous Data: Many State and Local Air Quality Agencies are currently busy examining PM-2.5 data to evaluate NAAQS compliance and to develop SIPs. States that have attainment problems for PM-2.5 will have to enact control strategies that tend to work by reducing one or more species or components of PM-2.5. These States will have to speciate their PM-2.5 filters as well as their PM-10 filters in order to know the effectiveness of these control strategies. PMc by difference provides filters useful for this purpose. These same States will need continuous PMc data to input into computer models to demonstrate how their control strategies will effect ambient concentrations far into the future.

Where and how much PMc sampling is needed?: It is apparent from AQS data in review document #4 “An Overview of PMc” as well as data collected by the New York State Department of Environmental Conservation (NYSDEC) that PMc mass is not necessarily related to population density and is only problematic in a few Regions. For these reasons, basing the requirement for PMc monitoring on population density such as was done with PM-2.5 would be a mistake. PMc concentrations are higher and more uncertain in areas affected by industrial and crustal sources. This places the majority of the need for PMc monitoring in smaller industrial cities and in areas impacted by wind blown crustal materials.

How strict should a standard be?: It is clear that many of the epidemiology studies to date that have focused on PMc data have used data suffering from poor bias and accuracy. I would prefer to see the current emphasis on producing accurate PMc data so that new epidemiology studies will be more robust in their determinations of causal effects. Until these new more accurate studies are performed, there is little justification to enforce an overly protective new mass based PMc standard. It is quite likely that for PMc, the epidemiology studies may find a particular species of PMc that needs to be regulated as opposed to the mass of PMc.

Why stick with the current PMc size fraction?: The proposed PMc size fraction of PM-2.5 through PM-10 dovetails with the NAMS, NCORE, STN PM-2.5 monitoring programs and the NAMS and TTN PM-10 monitoring programs. The future of EPA supported monitoring most likely is going to be guided by the NCORE program currently under development. One of the tenants of NCORE is the movement away from single pollutant networks to “coordinated, highly leveraged multi-pollutant networks”. The data from a future PMc monitoring program will be much more valuable if it can be used in a consistent manner with these other existing monitoring programs. It would be reasonable to switch monitoring size fractions in the future only if there was overwhelming epidemiological evidence in support of a different size fraction.

FRM, FEM, Does it matter? Do we need both?: As mentioned earlier, many Regions will need the ability to collect both filter PMc for speciation as well as continuous PMc for modeling. The best way to provide for all data needs is to approve a monitoring principle as the FRM for PMc. It is preferably that this principle be based on a technology that is not linked to a specific vendor so that future equipment development is not hampered. It is also preferable that the requirements necessary to obtain FEM status be such that the EPA will be able to support the use of one or more of the automated PMc monitoring technologies. Due to procedural rules in some States, it would be difficult for some Agencies to specify monitoring equipment that is designated as a FRM or FEM.

Difference Method

Question 1: The difference method is the best choice for providing the basis for approval of candidate reference and equivalent methods. It is the only method that uses a “fundamental measurement principle” to determine PMc. Since this method is weight based, it is consistent with the PM-2.5 FRM the PM-10 FRM and it works reasonably well in all geographic areas. The use of virtual impactors or optical/beta attenuation techniques by the other reviewed methods would include particulate properties in the PMc measurement that are not uniform from one geographic area to another.

Question 2: The accuracy and consistency of the difference method makes the data robust enough to compare to potential standards, use in health studies, provide filters which can be used for species analysis and for 24-Hr based modeling. Monitoring agencies are familiar with the field samplers and analysis issues and may in fact have extra samplers if the requirements for PM-2.5 FRM sampling are reduced. PMc network implementation costs would be reduced assuming that PMc sites were collocated with PM-2.5 FRM monitoring sites.

The disadvantages of the difference method are the limitations of a 24-Hr sample period, the delay in obtaining data, the additional uncertainty due to the operation of two samplers and the costs associated with field operations and remote lab preparation and analysis. Diurnal data is not available from this method though in areas where the PM-2.5/PMc ratio is fairly stable, diurnal information can be inferred from the hourly PM-2.5 network.

Dicot Filter Sampler

Question 1: The principle disadvantage of the filter based Dicot sampler’s use as a reference sampler is the deposition of a portion of the fine fraction on the coarse filter. This can be accounted for mathematically but the correction may not work well in all regions and it makes the method one step removed from a “fundamental measurement principle”.

Question 2: This method has the potential to be accurate enough for comparison with proposed standards and can provide filters for speciation analysis.

The disadvantages of the Dicot filter sampling method are similar to those of the difference method: the limitations of a 24-Hr sample period, the delay in obtaining data and the costs associated with field operations and remote lab preparation and analysis. Additionally, measurement problems associated with capture of the coarse fraction may make this technique less accurate in areas of the country with larger PM_c/PM-2.5 ratios.

Dicot Beta Gauge

Question 1: The Dicot Beta Gauge has not demonstrated consistent results from one geographic area to another. This shortcoming demonstrated by sub par comparison to the filter PM-2.5 and PM-10 data makes this a poor choice as a reference method or technique. Additionally, the lack of volumetric flow control will affect the instrument's performance in cold or high altitude areas.

Question 2: The only significant advantage to this method would be the availability of short term/hourly PM_c data and the reduced labor costs associated with an automated method. It is likely that this method could produce PM_c data useful for public health information but it would have to be periodically evaluated and adjusted against another method such as the filter difference method.

Disadvantages of the Dicot Beta gauge stem from fine particle intrusion onto the coarse filter, the unquantifiable effect of heating the sample stream and the extra error associated with the use of two Beta sources and detectors. One question that arises from the evaluation report is how exactly is the "span calibration performed by the user" related to the aerosol mass collected on the filter.

Coarse TEOM Method

Question 1: This technique has the greatest potential of the automated methods to be a reference or equivalent method. The instrument uses a single weight based measurement system and therefore does not calculate particle mass from a regional or aerosol dependant multiplier.

Question 2: The advantage of this design is the high flow rate virtual impactor which may in fact eliminate the issue of fine particle intrusion into the coarse measurement. This design element along with the question of the actual inlet cutpoint will require further evaluation. This method could provide hourly data useful for public reporting, trends and modeling analysis.

The principle disadvantages of this method include the heated sample, the lack of a fine particle measurement, and the inconsistent comparisons with the FRM difference data. It is

apparent that during the evaluation in Gary IN. on days like 4, 12, 25-27 the TEOM method drastically under predicts the data from the FRM difference method. It would be important to know if these days were associated with high PM-10/PM-2.5 ratios or high concentrations of volatile species.

APS Instrument

Question 1: The APS method would be a poor choice for a reference or equivalent method because it uses an assigned particle density to calculate a mass per unit volume. Average particle densities could be provided by regional evaluations but this can not effectively capture the short term changes in the density of aerosols particularly in urban and near source monitoring environments.

Question 2: The advantage of the APS method is its real-time data availability.

The disadvantages of the system are the lack of particle mass information, the inability to detect particles below 0.7 micrometers and the inconsistent comparison with the FRM difference method. The large differences between the collocated APS units in Gary IN. on sample days 10, 11, and 12 should be investigated.

Question 3: For the PMc DQOs, is the process the Agency took to develop the estimates of uncertainty appropriate? Are there factors the Agency has included that should not be considered or are there other inputs that should be included?

The process used to produce the PM Coarse DQO may or may not be appropriate but I worry about the unstated Type 3 error. The DQO tool was developed backwards, that is by examining the existing data in AQS. To actually know what the uncertainty is in any method you would have to study it from the beginning. You have to calculate the error associated with each step in a method from the preparation of the filter through the final data manipulation and presentation and then look at the propagation of those errors. For the PMc data used in this document, the potential errors are compounded by the use of vastly different and inconsistent sampling methods for both PM-2.5 and PM-10. Type 3 error is the problems that result from using inconsistent data.

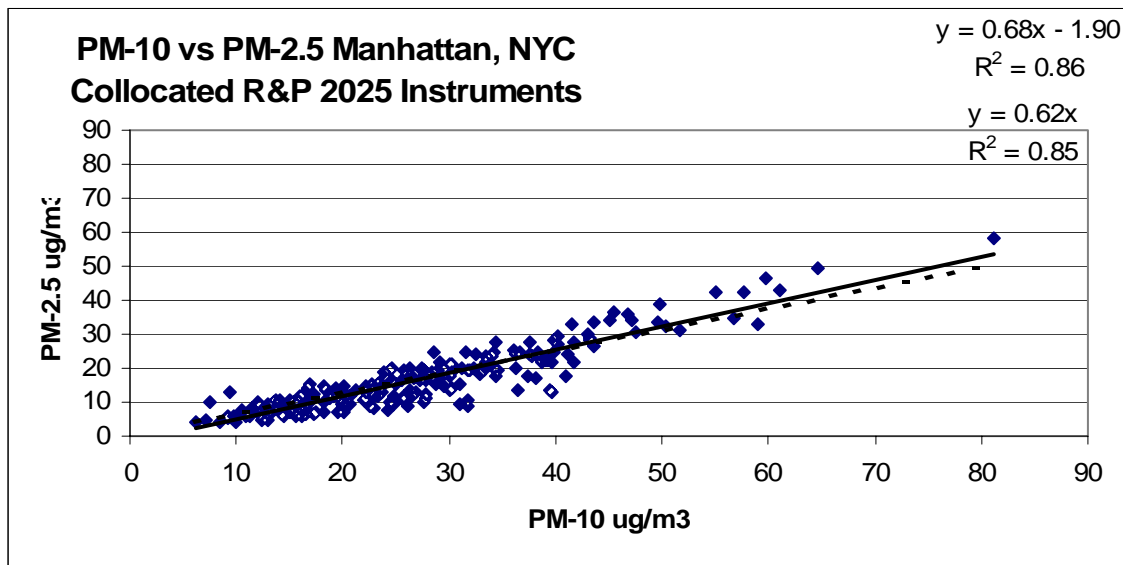
Inconsistent data should not be used comparatively. For example, low volume FRM PM-2.5 data includes a proportion of volatile material in the sample weight. High volume PM-10 samplers are not designed to capture this material and the resulting subtraction often creates a negative value for PMc. Similar inconsistencies result when using TEOM PM-10 data which is collected at 50 Deg C and is then compared to a different PM-2.5 collection method.

The DQO tool must be evaluated by using data that is of the same type and quality as the data that will be used in the future PMc network. This is the only way of knowing for sure that the resulting DQO is appropriate for the measurement. It would be advisable to initiate the PMc

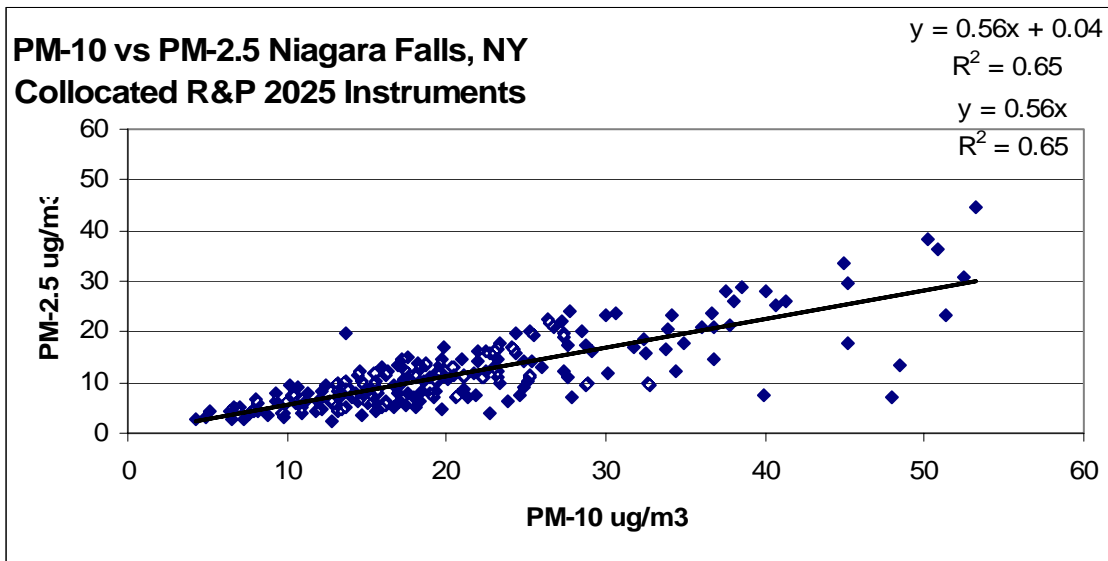
network with a start up phase that could provide data for the purpose of generating accurate DQOs.

Results of PMc measurements in NY by difference Jan. 2002 – May 2004

Manhattan	PM10 ug/m3	PM 2.5 ug/m3	PMc	PM2.5/PM-10
Average	25.40	15.47	9.92	0.60
Median	23.92	13.08	9.08	0.61
Std Dev	11.79	8.71	4.95	0.13
25th %	16.38	9.13	6.38	0.52
75th %	31.50	19.63	13.00	0.68



Niagara Falls	PM10 ug/m3	PM 2.5 ug/m3	PMc	PM2.5/PM-10
Average	20.25	11.42	8.84	0.57
Median	18.13	9.71	7.54	0.57
Std Dev	9.93	6.90	5.95	0.17
25th %	13.63	6.44	4.88	0.43
75th %	24.67	14.25	11.21	0.70



Dr. Rudolf Husar

Comments by Rudolf Husar, CAPITA, Washington University

Multi-site evaluations of candidate methodologies for determining coarse particulate matter (PM_c) concentration

EPA has conducted a set of field comparisons for candidate PM_c concentration measuring instruments. The project well conceived and executed and the report is well prepared. The Abstract could be more informative. Beyond listing the instruments and monitoring locations, it would be helpful to contain the key findings.

The field comparisons at Gary, IN (Mar-Apr, 2003); Phoenix, AZ (May-June, 2003); Riverside, CA (Jul-Aug, 2003) and Phoenix, AZ (Jan, 2004) have shown a remarkable precision for PM_c concentration measurements, which indicated that technologies currently exist for reliable PM_c measurements. However, the intercomparison field study has also revealed that substantial systematic deviations exist between some of the instruments, at some of the locations. The comments below pertain to these systematic deviations of PM_c concentration measurements.

The potential causes of systematic instrumental deviations can be related to either instrumental malfunction or to fundamental differences in the sampling/detection under different aerosol conditions. EPA, RTI and the instrument manufacturers have sufficiently addressed the instrumental malfunction issues. Therefore the comments below pertain to additional data analyses and characterization of the response characteristics of different instruments.

Additional analysis of existing TSI-APS size distribution data

During the field study the APS size distribution data were integrated to yield PM_c concentrations comparable to the other instruments. It would seem beneficial to evaluate the shape of the APS-derived size distributions for each integrated sample periods and analyzing the observed instrumental PM_c deviations in the context of the varying size distributions. For instance it could reveal that the magnitude of the deviations are related to the mass above 10 μm .

Furthermore, the APS data could also confirm or contradict the hypothesis of fine-coarse mass cross-contamination at the 2.5 μm size cut.

During the July 22, 2004 CASAC meeting the investigators have indicated the willingness to “make the APS data available to others”, but they did not present a plan for the re-analysis of the APS size distribution data. It is the opinion of this reviewer that virtually all of these instruments will be extremely sensitive to the shape of the size distribution at the upper cut-of 10 μm , which also influences the absolute PM_c mass determination. This is particularly significant since the upper end of the PM_c size spectra is highly variable due to the short residence time of ‘giant’ particles above 10 μm .

Additional instrumental inter-comparisons under controlled conditions

Field studies, as conducted during the first intercomparison study, are appropriate for initial evaluation of instrumental performances. However, field studies by themselves may not provide defensible explanations for the causes of the observed deviation. On the other hand, fully controlled laboratory comparisons, e.g. wind tunnel studies of instrument performance can be expensive and hard to control. A third alternative suggested herein is the ‘big room’ intercomparison, similar to the ‘big bag’ intercomparisons used for the study of short-lived ultrafine particles. In this approach, aerosols are generated, e.g. dust (by mechanical dispersion) or sea salt (by atomization) or any other substance is generated within a large pressurized room, such as a garage or hangar. The instruments could be mounted in the same configuration as during the field studies. Within the room, the aerosol would be circulated by large fans that assure the delivery of the same aerosol to all instruments. Since coarse and giant particles have relatively short life-time, during any given experiment the size distribution would change due to sedimentation of the giant fraction. Throughout the dynamic experiment the aerosol size distribution would be continually measured with continuous particle counters, such as ASP. Such an approach could yield exposure to partially controlled but fully characterized exposure to the instruments under a variety of aerosol conditions.

Sampling representativeness of PMc monitors

The design of PMc monitoring networks has to incorporate the issue of sampling representativeness. Sedimentation of giant particles above 10 μm limits their residence time to hours, minutes, or less. As a consequence of short atmospheric residence time, coarse and giant particles exhibit much stronger spatial-temporal variability than particles in the 0.1-5.0 μm size range. Thus, placing the monitors at the appropriate locations is crucial for obtaining the relevant PMc data. This location problem, of course, is hampered by the classical network layout paradox: the monitors can not be located optimally since the true emission/concentration fields are not known. If on the other hand, the concentration fields were known, there would be no need to perform extensive monitoring.

During the July 22, 2004 meeting several members have suggested the design and implementation of ‘saturation monitoring’ studies for purposes of characterizing the fine-scale spatial-temporal pattern of PMc at characteristic locations/seasons. It is the opinion of this reviewer that such studies would be necessary before the large-scale deployment of the new PMc sampling network. However, such intensive monitoring pilot studies would need to be augmented with significant analysis of the collected monitoring data. One particular crucial analysis would need to relate the temporal variability of PMc at specific monitoring sites to the spatial variability among the sites. Establishing such a space-time variability relationship from the saturation monitoring would allow the interpretation of continuous but sparse monitoring data to other locations. For instance, any continuous monitoring site that shows a relatively smooth background concentration and concentration spikes superimposed on this background provide clear evidence that a local PMc source is impacting that site. The interpretation of such monitoring data requires local wind data as well as a suitable mass conserving dispersion model for the evaluation of space-time relationship. The numerical analysis techniques for space-time analysis could be tested on the existing continuous PM_{2.5} monitoring data.

Estimating Parameters for the PMc DQO Tool

The PMc DQO estimation tool is a useful addition to the analytical toolbox of Regional, State and Local agencies. The report is well prepared. The comments below pertain to the formulation and testing of the model used in the PMc DQO tool.

Conduct model validation

The DQO tool is based on a statistical model that is formulated to describe the aerosol distribution pattern for estimating possible non-compliance. From science perspective, the statistical model need to fit a multi-dimensional PMc data space, including spatial dimensions (X, Y), temporal dimension (T) and particle size (D). The current DEQ model includes parameters for sinusoidal seasonality, phase shift between PM₂ and PMc seasonal peaks, day-to-day variability of PM₂₅ and PMc, the mean PM₂₅/PMc ratio, correlation between PM₂₅ and PMc.

In developing this empirical model, the DQO tool team has undoubtedly tested the validity of the above model for different aerosol conditions. It is therefore peculiar that the report is void of any information on model performance for ‘characteristic’ conditions. Was the model-data fit that poor? Where does it work where it does not? How can one use and trust a model (even for estimation purposes) if it has not been validated?

Alternative model formulations

The sinusoidally seasonal model with random daily perturbation and fixed PM₂₅/PMc ratio is just one possible statistical model of the aerosol system. Alternative formulations, such as non-sinusoidal seasonality; regional baseline + local (additive not multiplicative) perturbation; independent PM₂₅/PMc time series, would help establishing the robustness of the model(s).

Spatial aerosol parameters

The current statistical model simulates the temporal aspects of the aerosol variations. What about the spatial aspects? The addition of spatial parameters, such as spatial scale (e.g. derived from spatial cross correlation) for regional and local PMc would help estimating the spatial representativeness of different sites. Some of the spatial parameters could also be derived from the temporal monitoring data as discussed under the above heading: Sampling representativeness of PMc monitors.

Dr. Kazuhiko Ito

Comment on PMc monitoring documents. (Revised 8/01/04; original comment was submitted on 7/21/04).

Kazuhiko Ito, NYU.

(1) Multi-Site Evaluation of Candidate Methodologies for Determining Coarse Particulate Matter (PMc) Concentrations

Issues about the difference method:

In my original written comments, I expressed concerns regarding the use of the difference method ($PMc = PM_{10} - PM_{2.5}$), in particular the possibility of negative values. However, during the 7/22/04 meeting, at least two people, including one speaker, mentioned that this was not an issue with the current FRM samplers. The negative values apparently can be a problem when the PMc is computed based on the values from a “high-vol” PM_{10} sampler with a low-vol $PM_{2.5}$ sampler, but this is apparently not the case with the current FRM samplers (my original concern came from the data distribution of the computed PMc in Appendix 4, which included such high-vol PM_{10} samplers). Though I have not seen a database to support this point, my concern is dissipated after hearing from these experienced researchers who are close to these data. With this negative value problem being no longer an issue, and with the very high correlation of the data from co-located samplers in these data (R^2 is mostly > 0.95), my concern about PMc is now shifted to the spatial/temporal “error”, not the instrumental or analytical measurement error. I’ll discuss this in the second part of my comments (“The use of a performance based...”).

The great precision:

Since most of the discrepancies (mainly constant over- or under-estimation) between the samplers could be explained away and therefore some adjustments could fix them (or at least I got that impression), the great precision and the high correlation in most of these tests stood out.

I was impressed by the data, but I could also imagine that considerable care and efforts must have gone into these experiments, downside of which may be that, under more routine conditions, the same extent

Comments about specific candidate samplers:

I imagine that each of the candidate samplers must have an experimentally obtained collection efficiency curve. Showing such curves (and combined with estimated size distributions of PM in each location) would have been helpful in understanding some of the discrepancies across the samplers. Some of the reasoning for the differences between samplers remain unclear.

R & P sequential dichotomous sampler:

Unlike the difference method, the dichotomous sampler should not result in negative PMc mass. Therefore, the sequential dichotomous sampler seems more desirable than the difference method in estimating two mass fractions. It would be interesting to compare the dichot sampler vs. the difference method from FRM samplers in locations where PM_{2.5} dominates PM₁₀. The results from these high PMc locations look promising in terms of correlation with the FRM ($R^2 > 0.97$). The PMc under-estimation problem (due to particle loss) seems to be resolved.

R & P Coarse TEOM sampler:

Considering the major advantage of being able to provide real-time measurements, the high correlations between these TEOM samplers and the FRM samplers seem promising. If the consistent under-estimation of the TEOM sampler is in fact due to its smaller cut-off size (~9µm), then the unit may be re-designed to fix this problem. The large intercept (12.8 µg/m³) for the Phoenix 2003 data is somewhat worrisome, and the “better agreement” (apparently pointing to the mean TEOM/FRM ratio of 1.05) may be misleading if the ratio is systematically high at the lower PMc range and low at the higher PMc range. A large intercept was not observed in the Phoenix 2004 data, but PM_{2.5}/PM₁₀ ratios were also higher (0.24 for run 1-11 and 0.53 for run 12-15) during that period than the 2003 study period (0.18). As the time-series plot of the

TEOM and FRM from Gary, IN also suggests, the absolute difference appears larger at the higher PMc range.

Dichotomous beta gauge sampler:

A real time dichotomous sampler's strength is its potential ability to characterize diurnal patterns of fine and coarse particles, which may be also useful in separating out their corresponding source types. It is interesting that this sampler measures PMc more accurately than PM_{2.5}. While the report speculates that the over-estimation of PM_{2.5} is due to the inadvertent intrusion of coarse particles into the sampler's fine mode channel, it also mentions that PM₁₀ is also over-estimated (and PMc is not under-estimated in 3 out of 4 locations). The source of PM_{2.5} over-estimation needs to be clarified.

Aerodynamic Particle Sizer:

The obvious potential strength of this sampler is its ability to obtain size distribution of particles (larger than > 0.7 μm) in real time. A potentially problematic feature of this sampler is that it assumes a constant density for the coarse particles to obtain mass. It is not clear if this implies that site-specific determination (calibration) of coarse particle density is necessary. The extent of under-estimation of PMc in these data is rather large (~ a factor of two).

(2) "Use of a Performance Based Approach to Determine Data Quality Needs for the PM-Coarse (PMc) Standard".

The general rationale for the use of DQO process seems reasonable. It may be helpful if the reviewers actually get to try out this software with some scenarios.

There is one type of uncertainty that the document does not specifically address: spatial and temporal correlation of PMc within the scale of a city. From a viewpoint of conducting a short-term epidemiological study, the location of a PM monitor can be very important because the measurements at the monitor is supposed to represent the population exposure of that city. If the temporal correlations of PMc across locations within a city are low, then the associations between the health outcome and the PMc measured at such locations are likely biased toward

null (the difference in the absolute mean levels, not temporal correlation, affects long-term epidemiological studies). If the results from these epidemiological studies influence the setting of NAAQS, then the location-related uncertainty within the city should also be considered in the process of the DQOs.

The extent of this sampling location-related uncertainty for PMc, I imagine, would be far larger than the extent of instrumental bias and precision of co-located samplers reported in the multi-site field evaluation study. For example, Wilson and Suh (*J. Air & Waste Manage. Assoc.* 1997; 47: 1238-1249) examined site-to-site correlation of PM₁₀, PM_{2.5}, and PM_{10-2.5} in Philadelphia and St. Louis, and found that site-to-site correlation coefficients for PM_{2.5} were high ($r \sim 0.9$), but low for PM_{10-2.5} ($r \sim 0.4$; or R^2 of 0.16!), indicating that coarse particles have much larger errors in representing community-wide exposures. Compare this extent of correlation with those reported for the co-located samplers in the multi-site field evaluation study (R^2 s were mostly above 0.95). The uncertainty related to the location of PMc monitor can overwhelm the uncertainties related to instrumental measurement error. Furthermore, the location related uncertainty is also expected to vary regionally. Attachment 4 (“General characterization of PMc as found in the U.S....”) does not provide this information (site-to-site correlation within a city), but this can be computed from the same database for the cities where multiple monitors exist. At the 7/22/04 meeting, someone mentioned that the database in Attachment 4 contains both the “high-vol” and “low-vol” PM₁₀ samplers, so that the computed PMc in this database may have significant number of negative values (in fact, this is apparently the case as the 25th percentile of the Box plots are near zero in some areas in winter). Despite this limitation, I think it would be useful to characterize the spatial variability of PMc in the existing database unless there is a good reason to believe that the negative value problem may be regional and possibly blur the true spatial pattern of PMc.

Under the sources of uncertainty listed (the method, the NAAQS, the sample population, or the measurement uncertainty), the sampling location-related uncertainty of PMc would be categorized under the “Uncertainty Related to Sample Population”. However, the items listed here (seasonality ratio, population variability, auto-correlation, PMc/PM_{2.5} ratio, and PMc/PM_{2.5} correlation) are only indirectly related to the location related uncertainty. For example, two sites

can have an identical PMc CV (population variability coefficient of variation) and at the same time still have a very low temporal correlation with each other. This information needs to be somehow incorporated into the DQO process.

As I briefly mentioned at the 7/22/04 meeting, if the DQO model is to accommodate the input from epidemiological studies, then it will need to take into consideration two types of spatial/temporal variability: (1) temporal correlation across sites; and, (2) absolute difference in the long-term means across sites. Basically, the former is important for the short-term epidemiological studies (i.e., longitudinal and time-series studies) whereas the latter affects the long-term studies (i.e., cohort studies) in which the comparison is more cross-sectional. However, there are some further complications in defining the “error” or spatial/temporal variations. For example, in a simplest scenario for short-term studies, the PMc measurements made at each monitoring site within a city may be assumed to have random error that perturbs “true” citywide temporal fluctuations of PMc that affect the whole population. In such a case, the random error would attenuate (bias toward null effect) the true associations between PMc and health outcomes, but knowing the extent of the error (from the site-to-site temporal correlation) may allow “correction” of such attenuations. A more complicated scenario is that the PMc levels measured across multiple sites may not temporally correlated at all but nevertheless represent true exposures of the local population surrounding the monitor in each area. In such a case, monitoring at multiple locations may be necessary. A reality may be somewhere in between these scenarios. Likewise, for the purpose of long-term epidemiological studies, the variations in the long-term (e.g., annual) mean levels across monitors within a city may reflect both the true difference in exposure as well as some “error” in representing the population exposure for that city. The DQO model will need to somehow address these complications, but I think the first thing that has to happen is to estimate the spatial/temporal variation of PMc using the existing database. If the PMc based on the difference method for the older (high-vol) PM₁₀ data is problematic, then evaluating the spatial/temporal variation of PM₁₀ and PM_{2.5} separately may still provide useful information.

Dr. Donna Kenski

Comments on PM Coarse Methods Evaluation

Submitted to CASAC

by Donna Kenski

Lake Michigan Air Directors Consortium

2250 E. Devon Ave.

Des Plaines, IL 60018

July 30, 2004

In general, the documentation provided for review was quite thorough and the care with which the evaluation study was designed was evident. Most of these comments are minor, expressing a need for additional clarification or supporting material. It is tempting to pose some questions to EPA about the rationale for a PM_c standard and the evidence for health effects associated with this particular size fraction of PM, but that seems outside the scope of this subcommittee's charge. I trust that will be covered in detail in the coming criteria document.

It is obvious that no single method can meet all the goals for a PM_c monitor that were noted in the Scheffe memo of June 18. Thus a future PM_c network will likely consist of a mix of instruments to meet these various goals, much like the current PM_{2.5} network. The committee's charge to determine a suitable method for reference or a measurement principle is constrained by the definition of PM_c as the difference between PM₁₀ and PM_{2.5}; since PM_{2.5} and PM₁₀ already have reference methods, any PM_c method logically seems to need to be linked to these existing FRMs.

Simply by looking at how this current study was set up, it seems almost as if EPA has already decided, a priori, that the difference method based on a PM_{2.5} FRM with and without the WINS impactor is the reference method of choice for PM_c. Based on the data presented in Attachment 2, the difference method using these two instruments is probably the best understood, least biased, most consistent method available at the moment. However, this method is not without problems, and I have a strong preference for allowing EPA and the states as much flexibility as possible in implementing any monitoring, to avoid being locked into a single monitor design for the foreseeable future. Perhaps this flexibility is best built into the DQO process and into designating FEMs rather than FRMs. Nevertheless, the information presented in this study is not sufficient to make a definitive case for any of the methods at this time: additional monitoring methods should be evaluated (DRUM samplers), the manufacturers involved in this study have planned modifications that should be evaluated prior to any decision, and the data from this study should be mined comprehensively (i.e., using speciation analysis of all of the remaining samples, measured size distribution data, and inlet characteristics to more comprehensively examine sampler differences). As this study shows, each of the instruments had performance issues that need to be more completely understood before being considered as a reference method or measurement principle. These issues are summarized in the table below.

In evaluating the strengths and weaknesses of the candidate methods, I would like to have seen the raw data for each site in scatterplots, and consistent presentation of data from site to site. The

data presented in tables are useful summaries, but it's much easier to compare results graphically than in tabular format. For example, for the dichots, show the same set of scatterplots for each city: $PM_{2.5,dichot}$ vs. $PM_{2.5,frm}$, $PM_{10,dichot}$ vs. $PM_{10,frm}$, and PMc_{dichot} vs. PMc_{frm} , with a 1:1 line and a regression line on each plot. Subtlety and nuances in the data are lost when reduced to a single regression equation. There are two Tables labeled Table 6, and neither of them contains CV values; these should be added, since they are discussed in the text. As long as these tables are trying to summarize all the relevant information, the average measured mass for each species should be included as well, since it is an important factor to consider when examining relative performance of the various methods.

One drawback to the study was that, by virtue of being a research endeavor, it did not test the instruments under true "real world" conditions – for example, using local site operators instead of research scientists, using automated filter changing instead of running instruments manually (the sequential dichot), and weighing filters daily. The result is a more complete data set for evaluation (a good thing) but a biased view of the true field performance of the instruments. Comments on the actual ease of operation would have been helpful, although not necessary for evaluating performance.

Several issues were not addressed that really need to be considered. Chief among these is the presence of volatile material in the PM and the impact of that on these various measurements. Several methods employ a heated air stream; the effect of this on PMc needs to be documented. Similarly, the effect of the various inlets and their respective cut points should be examined. Allusions are made to inlet effects (re coarse particle intrusion onto $PM_{2.5}$, for example), but supporting data are not included.

Although the APS collected size distribution data, none of it was presented in the review materials. It would be informative to examine PMc size distribution characteristics.

Since there were several continuous methods, it would have been nice to see a comparison of diurnal data to evaluate how consistent these continuous methods are to each other, without regard to the FRM. This kind of comparison would be helpful in examining the various instruments' responses to changes in relative humidity and might be useful in evaluating their differences from the FRM as well.

Some additional background data that would be helpful in evaluating the methods include a description of spatial variability of PMc . Attachment 4 was a good start, but a quantitative estimate of the scale of PMc variability across urban areas, across states, and across regions would help. Similarly, since most existing PMc data (outside of this study) has been developed from older, high-volume PM_{10} measurements, it would be helpful to describe the comparability of the high-vol PM_{10} measurements to the low-volume measurements. A summary of PMc composition, or a conceptual model of PMc , would also add perspective.

DQO Tool: The process to date in developing the PMc DQO is appropriate, although the documentation is not the easiest report to read through. A question about Eqs. 1 and 2 in Attachment 3, Appendix A: the model assumes that $PM_{2.5}$ has a single distinct peak in each year, but in the Midwest it tends to have a bimodal behavior with two peaks (summer and

winter). What is the implication of this? Doesn't this mean the parameters (specifically for seasonality) incorporated in the model won't accurately represent data in a large part of the country?

Sampler	Advantages	Disadvantages	Comment
FRMc	Hundreds of similar samplers are already in use; field tested. Filter artifacts reasonably well understood.	24hr integrated measurement is less useful for public information goal, doesn't address diurnal variation, filter collection and weighing is resource-intensive.	Seems like the only really viable candidate for NAAQS, but must be supplemented by hourly measurements to supply the public with real-time information (if deemed necessary)
R&P sequential dichot	One instrument vs. two.	Consistently overmeasured PM2.5, undermeasured PMc, due to problems with sequential operation, but mfr. is addressing problem. 24hr integrated measurement is less useful for public information goal, doesn't address diurnal variation, filter collection and weighing is resource-intensive.	What fraction of the PM2.5 mass really ends up on the PMc filter? Should have some analysis of actual vs. theoretical.
Tisch dichotomous beta gauge	Semicontinuous method gives real-time data for public, usually cheaper/easier to run	1. Mass flow control will introduce errors when conditions change from calibration conditions. Such changes are inevitable in parts of the country with wide daily temperature swings. How big an effect? 2. Consistent overestimate of PM2.5 and PM10.;	Is volumetric flow control a possibility? Should address possible effects due to heating air stream to 25C. Should show how well theoretical adjustment for PM2.5 mass contained within PMc mass fits with actual. CV values are missing from table and report.
R&P Coarse TEOM	Semicontinuous method gives real-time data for public, usually cheaper/easier to run	Inconsistent: underestimates PMc in 3 of 4 trials, overestimates in 1, not clear why. Poor performance in rain.	Should address possible effects due to heating air stream to 50C. Why are the Phoenix 2003 results so different? Is this a nitrate volatilization problem?
TSI APS	Semicontinuous method gives real-time data for public, usually cheaper/easier to run	Particle density (which is presumed to be 2 g/cm ³) will vary with site and particle composition. Need documentation to support this value and description of variability. Underestimate PMc. Correlation with FRM great in Phoenix, but poor in Gary and California—why?	

Dr. Thomas Lumley

Comments on material for July 22 meeting.

Thomas Lumley
Associate Professor of Biostatistics
University of Washington.

0. One initial issue is whether the definition of PM₁₀ is on the table for discussion at this point. The cutpoint at 10 microns is maximally difficult for precise and accurate measurement. If there are biological grounds for moving the cutpoint it might make it easier for different measurement methods and instruments to agree on PM₁₀ concentrations.

1. There are high correlations but relatively poor agreement in values between the measurement techniques compared in the EPA study. The ratios between the methods vary from site to site, implying that the instruments are not all measuring the same thing. I do not have the right expertise to comment on which of these things is the best for defining as PM₁₀, but it seems important to understand the reasons for differences between the measurements

- Do the size distributions from the APS suggest that the differences are due to different cutpoint characteristics at 10 microns? Size distributions would also be of interest in Phoenix for the times when 'intrusion of coarse particles into the fine mode' was postulated for the R&P dichot sampler.
- Are there chemical composition or other data that would illuminate the likely impact of semivolatiles on the differences?
- A more difficult question to answer is whether there are day-to-day or seasonal variations in the relationship between the measurement techniques (for example, in Phoenix, is the relationship different on the windier days?). The reason this is important is that a stable relationship would allow site-specific calibration of one method to another.
- The variation between sites in the ratio of APS to FRM measurements suggests that a single global density value may be inadequate. This in turn raises the question of whether a single density value is adequate across size categories within a site.
- Scatterplots comparing results of different measurement methods (like Figure 10 of the evaluation document) would be very helpful for other comparisons, especially if they included all the sites rather than just one.

Clearly the continuous-time methods have the ability to describe regular diurnal variation and brief episodes of high concentration. The APS instrument can in addition describe size distributions. Time-resolved (and size-resolved) information may be helpful in reporting and understanding brief episodes of elevated PM. On the other hand the continuous-time measurements do not collect PM in a form suitable for subsequent chemical speciation.

2. The performance curve tools described in Attachment 3 are a very promising development. When relating observed data to the true concentrations specified by the NAAQS there is unavoidable uncertainty due to the variability of PM concentrations and the bias and limited precision of measurement techniques. It is important to be able to relate the uncertainty in concentration to the probability of making an incorrect decision, and to decide whether the cost of reducing this probability is warranted.

The statistical model and simulations used to develop the curves are thoroughly developed and generally well-described. Some specific points:

- I am uncertain as to whether the autocorrelation parameter described is the autocorrelation of the log-Normal errors or of the Normal errors that are used to generate them. Either would be reasonable, but a user would need to know which was intended, as they will be quite different.
- Although not estimable from the AQS database, the correlation in measurement errors between PM_{2.5} and PM_c may affect the results, and may vary between measurement techniques. For example, computing PM_c as PM₁₀-PM_{2.5} introduces a negative correlation between errors in PM_c and those in PM_{2.5}. This correlation is likely only to be important when considering the two size fractions jointly (e.g., what is the probability of a location being incorrectly declared in compliance on both size fractions?)

3. One standard but perhaps undesirable feature of the current calculations is that they assume that the action limit, rather than the true concentration, is fixed. Considering a hypothetical PM_c standard of 50mcg/m³, for example, it would seem statistically natural to examine how different action limits perform in ensuring a true concentration below 50mcg/m³. The description in Attachment 3 reverses this, assuming that the action limit is prespecified and examining how the true concentration would in fact be controlled.

In addition to the fact that the true concentration is presumably the quantity relevant to public health, the interaction between cost and precision is easier to handle when the true concentration is used as the target. Consider two locations, one with very low PM_c levels, and one with moderately high levels, just below the permitted threshold. In the first location even relatively imprecise measurements will be sufficient to show that the PM_c levels are in compliance. In the second location much more precision is needed. If the monitoring specifications were designed to ensure, say, a 95% power for detecting a true level of 55 mcg/m³, the first location could use less precise measurements and a correspondingly lower action limit than the second location. In both cases the public health would be protected, but at lower cost than if both locations used the more precise monitoring.

The calculations for varying the action limit are almost the same as those for varying the true concentration, and similar curves (but sloping down rather than up) would be produced.

Dr. Peter McMurry

Peter H. McMurry

July 23, 2004

RE: CASAC Ambient Air Monitoring & Methods (AAMM) Subcommittee Meeting: Comments on Materials provided for review

Preliminary Comments:

I am impressed with the amount of work and thought that was done prior to this meeting. I am also delighted with the responsiveness of EPA to recommendations previously made by CASAC's Technical Subcommittee on Particle Monitoring. It is clear that input from this committee is having an impact on EPA's decision making process, and this is gratifying.

I am impressed (and surprised) by the excellent precision of measurements obtained with using identical samplers operating in parallel. My compliments to the manufacturers and the field measurement team on a job well done!

I agree with the accommodating tone of the report. For example, "Effective engineering solutions to this noted problem could potentially result in close agreement of the R&P dichot with the filter based FRM for all three metrics." (Summary, p. 23) Manufacturers *should* be given an opportunity to refine instrument performance.

The DQO tool is intriguing. However, I will restrict my comments to measurements methods, since this is my primary area of expertise.

Responses to Questions:

1 & 2. What are the strengths and weaknesses of each method tested in the ORD study

Method	Strengths	Weaknesses
PMc FRM=PM10-PM2.5	<ol style="list-style-type: none">1. Uses established monitoring equipment2. Filters can be analyzed for particle composition	<ol style="list-style-type: none">1. 24 hour time resolution2. Expensive, manual filter analysis3. Not useful for real-time AQI reporting4. Involves collection of particles on a filter, rather than direct measurements of gas-borne particles.5. Relies on FRM, which is known to be inaccurate in some locations

Dichot	<ol style="list-style-type: none"> 1. Single sampler for both fine and coarse mass/composition 2. Less influenced by fine/coarse missing than PM10-PM2.5 3. Filters can be analyzed for particle composition 	<ol style="list-style-type: none"> 1. 24 hour time resolution 2. Expensive, manual filter analysis 3. Not useful for real-time AQI reporting 4. Involves collection of particles on a filter, rather than direct measurements of gas-borne particles.
TEOM	<ol style="list-style-type: none"> 1. Fast time response-better information on temporal exposures 2. Established track record 3. Data can be used for real-time public reporting 4. May currently have greater potential for accuracy than other methods. 	<ol style="list-style-type: none"> 1. Volatilization losses for high temperature collection 2. Apparent sampling losses for coarse particles. 3. Involves collection of particles on a filter, rather than direct measurements of gas-borne particles.
Beta	<ol style="list-style-type: none"> 1. Fast time response-better information on temporal exposures 2. Established track record 3. Data can be used for real-time public reporting 	<ol style="list-style-type: none"> 1. Fine mass measurements are in poor agreement with those from other samplers. There must be a lot of FRM-beta gauge comparisons from previous fine particle measurements: are such results typical? 2. Involves collection of particles on a filter, rather than direct measurements of gas-borne particles.
APS	<ol style="list-style-type: none"> 1. Provides valuable supplemental information on size distributions. 2. In-situ measurements of gas-borne particles (the type we breathe!) 	<ol style="list-style-type: none"> 1. Not a mass measurement method; should not be considered as such

Recommendations:

1. Consider adding the DRUM sampler, developed by Cahill and coworkers, for future instrument evaluation studies. The DRUM can provide size and time resolved information on elemental composition, absorption, and mass (through beta attenuation). Impactors operate on a different sampling principle than those that have been studied to date. While all approaches have pros and cons, impaction offers the benefit that coarse and fine particles are not mixed together on the substrate, and that collected samples are less prone than filter samples to evaporative losses. I am not aware that a commercial prototype of this instrument is currently available, but I assume one could be developed quickly if the instrument were found to perform well.
2. Future atmospheric measurements should be carried out in airsheds that would provide new challenges to coarse particle measurements. Examples of such phenomena include high concentrations of coarse biological particles (e.g., pollens), the presence of volatile compounds (e.g., organics, ammonium nitrate), etc.
3. I am of the view that the APS would not be an acceptable measurement instrument for coarse mass compliance measurements, since it does not measure mass. Measurements are complicated by variabilities in particle density and shape factor. These particle properties are likely to vary spatially, temporally, and among particles of a given size at a given instant of time. I do not think it will be possible to reduce to an acceptable level uncertainties in estimated masses obtained with this instrument. On the other hand, the APS would be an excellent instrument for intensive field studies of coarse particle size distributions, or of short-term temporal and spatial variabilities coarse particles concentrations.
4. If measurements of spatial and temporal variabilities were to be carried out using an array of APS instruments, consideration should be given to measuring the complete size spectrum (3 nm-10 μm) at the same time. There is interest in spatial and temporal variabilities of ultrafine particles, for example, and the sampling methodology and expertise required to study this question would be the same as that required for studying coarse particles. The Supersite program has led to the development of instrumentation systems and skilled personnel who could carry out such measurements.
5. I feel that EPA should require that all size-dependent efficiencies be characterized for any instrument that they approve for compliance measurements of course particles. These include size-dependent sampling inlet efficiencies, and size dependent collection efficiencies and deposition losses in virtual impactors. In the latter case, measurements should be done for both solid and liquid particles. Such information will provide valuable insights into measurements obtained in different environments. Careful measurements of sampling efficiencies for the standard PM₁₀ inlet have been carried out. The modified PM₁₀ inlet which is being used for the high flow TEOM has not been studied with the same amount of care.
6. The report suggests that differences between the R&P PMc TEOM and PM FRM may be due to the 9.0 - 9.5 μm cut point of the TEOMS' inlet. An effort should be made to determine whether or not the measured size distributions substantiate this hypothesis.

7. Can the measured size distributions provide insight into the extent to which coarse particles may have intruded into the R&P dichot fine channel? This could be calculated from measurements by making use of size-dependent efficiency curve for the R&P inlet (assuming that they are known.)

8. I think it might be desirable to develop a PMc measurement technique that does not make use of the PM2.5 FRM measurements. This would avoid confounding the PMc measurements by fine particle sampling artifacts, which are known to occur in some environments. It would also open the possibility to a methodology that provides automated measurements with higher time resolution. Such measurements would provide data that would be useful for a wider range of purposes than simply compliance monitoring.

9. I think future sampler characterization research should focus *primarily* on atmospheric sampling as opposed to laboratory studies . The exception is measurements of size-resolved sampling and collection efficiencies mentioned in point 5 above.

Dr. Kimberly Prather

U.S. EPA/CASAC/AAMM Subcommittee

Kimberly Prather

Comments after the July 22, 2004 Meeting

The subject of a new coarse particle standard was discussed. We were informed the standard would be PM_c (PM₁₀-PM_{2.5}). The goals of the meeting were to help EPA identify a strategy to allow them to choose the best measurement technologies for monitoring PM_c.

First of all, the studies performed by EPA were clearly well thought out and conducted. I have several comments with regards to the PM_c standard as well as possible future studies:

- 1) The addition of the APS at all sites was a nice one as it could provide insight into changes in size distributions. From the discussions, it sounds like it was chosen to help understand observed differences between measurements and not to measure PM mass (although that was investigated). I agree it would be useful to try and use these data to help understand the observed differences. It seems like more could have been learned if the PM₁₀ inlets had not been used (as one committee member stated, you can always impose that cut-point later on) so you can directly measure how shifting of fine and coarse size modes affects all measurements.
- 2) Differences in composition at all sites need to be better understood so the observed differences in methods can be addressed. For the next round of studies, it would be worthwhile for EPA to add semi-continuous measurements of nitrate, sulfate, and carbon using commercially available instruments (rather than completely relying on filter-based sampling). It is likely the instrument manufacturers would “loan” EPA the necessary instruments for these studies.
- 3) The issue of why/how PM_c was chosen came up at the meeting. It appears (based on the responses given) that the reasons for choosing this particular definition for the coarse standard are not well founded. There could be issues with further convoluting two problems. Oftentimes it is questioned as to why PM_{2.5} was chosen because it does not effectively separate the modes (or PM sources) well (i.e. PM₁ would be better at doing this). Now in choosing PM₁₀ as a cut-point, this does not pull out coarse mode particles well either (it is on the lower size shoulder of the coarse mode a large fraction of the time). It would be most appropriate to use established size distributions and source contributions to those distributions to establish the best cut-point. The committee was told just to focus on the instruments (for this meeting) and that we could come back to the PM_c standard issue another time. My feeling is once we spend all the time and effort choosing the best measurement methods and do further (costly) studies, no one will want to go back and re-visit this issue. Since EPA is being forced to re-visit the best PM standard, now would be the time to choose a cut-point that can be backed up by strong science. A comment was made that we have the most health data for PM₁₀. I would argue that since the available health data yield conflicting conclusions on the health impacts of PM₁₀, maybe this is telling us there is a problem with where the line is being drawn. The key issue is to choose a cut-point that “makes sense” based on the physical size distributions and source contributions to that

size distribution (both of these area pretty well established). The proper cut-point which separates sources will make understanding health impacts more straightforward and also allow establishment of proper control strategies.

4) It was not clear if EPA is trying to choose only one measurement technique for monitoring coarse PM or a suite of instruments. Ideally, one should choose a suite of instruments that yields information on short term variability, size distributions, and allows speciation to be performed. The trade-off of course is as the cost of the combination of instruments grows, less sites can be studied (yielding less data on the spatial variability of PM).

5) The report did not give meteorological data which would be important for linking short term variability with sources impacting an area. Additional knowledge on source impacts on the coarse particle standard will be important for developing appropriate control strategies for different areas.

6) One cannot underestimate the importance of performing some control (lab) studies to understand discrepancies observed in the field. By this, I do not mean using pure (unrealistic) particles, but using particles that could potentially contribute to PM_c and cause differences in the methods such as resuspended dust. In these studies, one could generate known aerosol compositions, concentrations, and sizes and introduce them to all instruments simultaneously. The comment was made that lab studies are not as good as field studies since it is difficult to simulate “real” particles. This is true, but I believe carefully conducted lab studies (complementary to those conducted in the field) using “real” particles that could give you the biggest discrepancies is important for isolating the issues contributing to the observed discrepancies. Oftentimes in the field, too many things are changing at once—lab studies allow you to isolate these changes and study their effects. The pre-warning is there are an infinite number of possible lab studies that could be conducted, so careful planning (based on field results) is important.

7) A key point in picking an instrument is deciding the benchmark which represents the “true” answer. It appears the method being used in this capacity is the FRM which is troublesome since it has known issues. It certainly has been used to obtain a wealth of data, but this is not a valid reason for continuing to use it. It only propagates the problem. Choosing a method because it agrees with (or is consistent with) another instrument that has problems is not the best approach. Time needs to be invested in seriously choosing a PM material that can be used (with the chosen instrument/s) that will allow one to ultimately operate the instrument/s with the best accuracy.

8) What will be the time resolution of the standard? I do not believe 24 hours is short enough—coarse particles (particularly those emitted from sources such as coal combustion) show rapid variability and very high values over short time periods. These high concentrations will be averaged out over 24 hours. Since we do not fully understand the health impacts of these particles, capturing these fluctuations is extremely important. Thus, choosing an instrument that does “real-time” measurements will be important.

9) It wasn't clear if all of the inlets were the same in the report. These inlets should be fully characterized before any further field studies are performed.

10) Semivolatiles could be important in some regions and possibly account for some of the observed differences (especially in changing the PM_{2.5} fraction). Some of the instruments used were heated, others were not. It would be worthwhile to study how these species affect the results. The toxicology of these compounds is pretty well established so their contributions should not be ignored.

Dr. Armistead (Ted) Russell

Comments on PMc Methods Evaluation and DQO reports – Ted Russell

Since I am getting these comments in a bit late in the game, I will try not to go over the type of input provided so far.

First, the EPA staff and their contractors should be congratulated for this piece of work, both the methods intercomparison and testing, and the DQO tool. A few things, though.

First, in the spirit of no good deed goes unpunished, I have been on the NAAMS Subcommittee, and the good work that EPA did as part of putting together the NAAMS, they should thoroughly consider and emphasize how the various methods fit in with that strategy, and ask the Subcommittee to address that issue as well. They do address this, though I would say more tangentially than directly, and their questions to the subcommittee do likewise.

Next, an overarching consideration, in part building upon Warren White's comments, but with my own concerns, is that the ordering of the questions suggests that determining a reference method is more important than achieving multiple monitoring objectives. From a regulatory standpoint that may be true, but from an overall air quality management perspective that is not true. Along those lines, I would not like to see us start emphasizing the use of an integrated sampling approach, providing 24 hour samples every 3-12 days, as opposed to a (semi) continuous sampler. Adoption of a 24-hour approach can inhibit the deployment of more informative methods. (On the other hand, the use of a 24 hour method that also includes speciation has some attractiveness.) As noted by White, the ability to identify correlations with other pollutants and atmospheric variables is important for understanding the problems, and thus the solutions. Having only 24 hour samples, taken every 3-12 days, greatly inhibits that ability. As a modeler, I should add that having a 24-hour mass measurement is not great for evaluating a model. If you are off a long ways, that is bad. If you are close, it is difficult to say if the reason is that you have the processes about correct, or if there are compensatory errors. Indeed, one can get the diurnal behavior backwards and still find good performance.

I see a second problem with an integrated sampler in that it presupposes the form of the standard, which is presumed, at this point, to take the form of an annual average limit and a 24 hour limit. First, for the annual standard, I suspect some of the (semi) continuous samplers can

provide very comparable results to the difference methods, so that is less of an issue. For example, the TEOMs from the evidence provided, and seeing their operation for PM_{2.5}, give very similar results averaged over that long of a period. Even over a one-day period, we see that they agree pretty well. But then we should ask, why a 24-hour standard? If exposure to high concentrations is the issue, 24 hours is probably not the period of interest: one to eight hours likely makes more sense. PM_c exposure will likely happen predominantly outdoors as many ventilation systems will remove this fraction of PM, and it has a more limited lifetime indoors. Thus, the likely exposure will be for more limited times. Integrated samplers using the difference methods are less able to capture the needed data for the shorter time periods for two reasons. First, unless one is committing to tremendous labor commitments, identifying the maximum eight hour average in one year is not likely. Second, the difference method gets less reliable as the masses on the filters become smaller.

Putting the two issues above, together, also lead to a third reason not to have an integrated sampler become the method of choice. If we continue to have, predominantly, 24 hour samples upon which to conduct health-related research, we will continue to find associations with 24-hour metrics, which would then support 24 hour standards. If we were to have more short term measurements, associations with other intervals might be found to have a stronger association, and thus argue for a different form of the standard. It should also be noted that PM_c has a shorter atmospheric lifetime than PM_{2.5}, so we would expect to see more diurnal variation, further suggesting the appropriateness of a shorter period for a standard, if, indeed, health effects are related to a shorter term exposure. (For a long-lived pollutant with relatively little diurnal variation, the difference between using a 24 hour or 8 hour standard has less impact than if it is a short-lived pollutant.).

With the above in mind, my answers to questions one and two as posed to the Subcommittee are intertwined, and I have to speak primarily as a modeler, not a instrumentation expert. Obviously, I would like to see the use of a continuous sampler for both. The data provided, and recognition that TEOMs enjoy relatively wide spread use in the field, suggest that they are an attractive method, either as a reference method or equivalent method, e.g., after undergoing the process outlined in the NAAMS draft. The APS is a non-starter, at present for a number of reasons in terms of the documented performance, as well as reasons given by others. The difference method appears to be very repeatable between samplers, even by different brands,

and may provide speciated data as well if one chooses (though this can be provided by adding speciation filters to the other systems as well).

Dr. Jay Turner

CASAC AAMM Evaluation of Coarse Particle Monitoring Methods Comments Submitted by Jay Turner on July 20, 2004 in Preparation for the July 22, 2004 Meeting

Per the June 18, 2004 memorandum from R.D. Scheffe to F. Butterfield, the Subcommittee has been charged with three questions: “(1) what are the strengths and weaknesses of each method tested in the ORD study for the purposes of using it as a reference method, a measurement principle, and as a method that would provide the basis for approval of candidate reference and equivalent methods; (2) what are the strengths and weaknesses of each method tested to meet multiple monitoring objectives such as comparison to potential PMc standards, public reporting, trends, chemical speciation, and characterization of short-term episodes and diurnal variation; and (3) for the PMc DQOs, is the process the Agency took to develop the estimates of uncertainty appropriate and are here factors the Agency has included that should not be considered or are there other inputs that should be included?” My written comments submitted prior to the July 22, 2004 meeting focus exclusively on the first two questions.

What are the strengths and weaknesses of each method tested in the ORD study for the purposes of using it as a reference method, a measurement principle, and as a method that would provide the basis for approval of candidate reference and equivalent methods? The question of a benchmark methodology¹ begs us to first consider our overarching objective(s), as a frame of reference is needed to assess strengths and weaknesses of the tested methods. If PM_{2.5} serves as a model, we cannot expect to accurately quantify PMc mass concentration by any single measurement. Therefore, we first need a conceptual model for PMc – its physical and chemical properties – and subsequently need to determine the salient features which we desire to capture (possibly at the expense of accepting bias in other features). For example, in the case of PM_{2.5} FRM methodology the role of aerosol bound water is addressed by conditioning filter samples at a fixed relative humidity; we are not measuring the true aerosol content (nor do we necessarily desire to do so in this case) but rather bring all samples to common conditions which is at least interpretable. In one light this might be considered a weakness (significant departures from reality) and in another light it might be considered a strength (typically suppressing the role of aerosol bound water and to a large extent harmonizing across the network). As another example, we know that the PM_{2.5} FRM methodology can be susceptible to large losses of ammonium nitrate; this is largely considered a weakness and thus another method that properly quantifies the ammonium nitrate might be favored if any corresponding tradeoffs are deemed acceptable. The relative roles of water, organics adsorption, nitrate losses, and so on are likely quite different for fine PM and coarse PM; are we in a position to articulate the various factors that might influence PMc measurement and prioritize the most important features which any benchmarking methodology should capture? I am reticent to move to quickly and deeply into casting the methods comparability in terms of strengths and weaknesses until we holistically (to the extent

¹ I intentionally lump the three stated uses (reference method, measurement principle, basis for approving reference and equivalent methods) into simply “benchmarking methodologies”, for the present time ignoring the important distinctions between the three uses.

practicable) describe the system we are trying to measure. That said, some comments are warranted based on the ORD study.

Integrated filter measurements offer simplicity at the expense of labor intensiveness and data incompleteness (the latter may or may not be an issue, depending on the outcomes from the DQO Process). The PM_{2.5} and PM₁₀ measurements in the ORD study feature high precision and thus the propagated uncertainty from the PM₁₀ minus PM_{2.5} differencing method for PM_c mass concentration might be acceptable. An underlying assumption – still to be verified – is that any artifacts in the fine fraction of PM₁₀ are the same in the separately collected PM_{2.5} sample (that is, there are assumed no matrix interactions between the fine and coarse particles on the PM₁₀ sample). The dichotomous sampler method has reduced sensitivity to the PM_{2.5} precision – a smaller adjustment is required to correct the coarse channel mass concentration for fine particle intrusion than to correct PM₁₀ for PM_{2.5} – but there are differences in the flow rates for the fine and coarse channels which may affect the relative role of fine PM artifacts between the channels and bias the applied correction. Other dichotomous sampler issues warranting attention include the virtual impactor design and losses in filter shipping for conditions typical of network deployment (rather than a special research study). In the latter case, the ORD study results are encouraging but more work is needed.

Semicontinuous measurements are less labor intensive (assuming the instruments are indeed field robust) and offer greater data completeness. The temporal averaging of near real-time measurements may suppress certain artifacts inherent in the temporal integration of the substrate sample collection. Within certain constraints, there are opportunities to tune the instrument design (e.g., inlet heaters) to address the most desired features. On the other hand, if one can use PM_{2.5} as a model it appears there might be no “one size fits all” approach to such tuning and again we must consider strengths and weaknesses in light of the to-be-defined features. The ORD study presents the semicontinuous methods to the filter methods (specifically, the differencing method). If the goal is a semicontinuous method which quantitatively aggress with the differencing method, the preliminary results are encouraging but there is much more work to be done to reconcile the differences and hopefully improve the comparisons. It is not clear to me at this point, however, that this is the explicitly-stated goal. Furthermore, the comparisons are clouded by differences in inlet cutpoints and sample stream conditioning (e.g., use of inlet heaters); it would be very helpful to have well-characterized inlets and a better understanding of how the conditioning might affect the respective measurements.

The ORD study provides a rich data set for probing key questions concerning the status of PM_c measurements. The first report emerging from that work - “Multi-Site Evaluations of Candidate Methodologies for Determining Coarse Particulate Matter (PM_c) Concentrations” - covers substantial ground towards presenting and interpreting the results. It is acknowledged that the report cannot be exhaustive, and the following specific comments are offered towards placing the presented results in context and suggesting items that could be probed in the existing data set.

- (1) It would be very helpful to see scatter plots in addition to times series for most of the comparisons provided in the report. Even in cases where the R² is high, there is still much to be learned from scatter plots in terms of the quality of the agreement between different methods.

- (2) While the measurement methods evaluation should not presume a specific form for a PMc standard (e.g., averaging time and threshold concentration), historical approaches to regulating PM suggest that both long-term averages (quarterly, annual) and relatively short-term averages (e.g., daily) might be considered. The first ORD report focuses largely on study-average results (e.g., mean ratios)²; it would be very helpful to show how the instruments compare for conditions that would be representative of possible violations of a short-term averaging period. For example, is there any degradation in the collocated precision at high daily-average concentrations? Is there any degradation in the agreement between the semicontinuous measurements and filter differencing measurements at high daily-average concentrations? The data set might be too small to address these issues in great detail, but some elaboration would be helpful (again, scatter plots would be a first step).
- (3) One stated possible explanation for the Coarse TEOM underestimation of PMc (relative to the FRMs) is the inlet cutpoint being closer to 9 μm rather than 10 μm . Can the APS data shed light on the extent to which the discrepancy can be reconciled by the difference in inlet cutpoints?
- (4) While the Tisch SPM-613D units seem to perform well for PMc, the large differences for PM_{2.5} are disconcerting. A clear understanding of the source of the discrepancy must be elucidated before the instrument could be used with confidence.

What are the strengths and weaknesses of each method tested to meet multiple monitoring objectives such as comparison to potential PMc standards, public reporting, trends, chemical speciation, and characterization of short-term episodes and diurnal variation? The aforementioned first report by ORD provides little insights into the data relevant to addressing this question (although there is likely substantial information which could be mined from the study); thus, my comments are presented in general terms. The filter methods are superior for chemical speciation and the semicontinuous methods are superior for public reporting and characterization of short-term episodes and diurnal variation. I make these statements with qualification, however, as I am inferring robust methods are being used. For chemical speciation, while the filter differencing method shows promise for having adequate precision in the PM_{2.5} and PM₁₀ mass concentration measurements such that the PMc mass concentration would have acceptable precision, it is not clear whether this would be the case for all chemical components of interest. An analysis is needed to determine whether the differencing method would be robust for chemical components and not just total mass concentration. The semicontinuous methods clearly are superior for public reporting if the goal is timely reporting; as for diurnal variations, the semicontinuous measurements are valuable if the measurement biases (which may have diurnal variation depending on their origin) do not mask the true diurnal structure. Both filter and semicontinuous methods may be suitable for comparison to the PMc standards although there are certainly tradeoffs and different forms of the standard may warrant different DQOs which in turn influence the selection of a preferred measurement method.

² Note that C.V. values were not reported in Tables 6 and 7.

In summary, the ORD study offers very valuable insights into measurement methods for PMc; with additional conceptual work and mining of the data we can move towards evaluating the suitability of these methods (in current form and subject to design/operation changes) and possibly other methods.

Dr. Warren H. White

Comments on PMc Methods Evaluation and DQO reports – Warren H. White, 7/10/04

I am happy with the homework EPA is doing on PMc measurements. The Agency is showing admirable initiative in this area, and the technical quibbles I raise on the following pages are intended mainly to reinforce the diligence it is already exhibiting.

An important concern for the methods evaluation and DQO process is that they not focus too narrowly on NAAQS attainment decisions. The charge to the subcommittee takes note of other monitoring objectives such as public reporting, trend detection, and episode characterization. I would like to highlight another that is at least as important as any of these, the accurate determination of coarse particles' covariance with other particle fractions and gas species. The characterization of this covariance requires measurement precision well beyond that needed to establish an annual mean, and at concentrations well below those of concern for a 24h standard.

PMc measurements should be as precise as feasible because (White, 1998, *J. Air & Waste Manage. Assoc.* 48, 454-458)

- (a) noise always depresses observed correlations with other measurements,
- (b) historical PMc measurements have been much noisier than associated PM_{2.5} measurements, and
- (c) crucial inferences about health effects and atmospheric behavior are routinely based on the differences observed between correlations involving PMc and PM_{2.5}.

(For a ready example of the last phenomenon, consider page 9-16 in the June 2004 draft Criteria Document: “new data reinforce our earlier understanding that ambient concentrations of fine particles (measured as PM_{2.5}) are typically more highly correlated and/or are more uniform across community monitors within an urban area than are coarse particles (measured as PM_{10-2.5}), ... Thus, central site ambient concentration measurements are a better surrogate for population exposure ...”)

1. The inter-manufacturer precisions given in Table 2 of Attachment 2 for the collocated FRM samplers in the evaluation study are surprisingly good. I gather that these numbers come from the formula in Appendix 3B of Attachment 3, which subtracts out the contribution resulting from the mean difference between measurements. Is that correct? If so, it would be useful also to tabulate the observed root-mean-square differences to give a more direct picture of the overall measurement uncertainty.

It is also worth noting that the high $PM_{10}/PM_{2.5}$ ratios sampled in the evaluation study maximize the precision of the $PM_{10} - PM_{2.5}$ difference measurement. They are not representative of conditions in many of the studies that find little association between PM_{10} concentrations and health effects. The six cities analysis of Schwartz et al. (1996), for example, was based on $PM_{10.2.5}/PM_{2.5}$ ratios that averaged less than 1/2.

2. The foundational discussion of Figure 1 in Attachment 3 could be much clearer. Defining the “gray zone” as the S-shaped region between the performance curves is, first of all, a distraction because the visual area of that region is only weakly related to the amount of indeterminacy facing the decision-maker.* The true gray zone for *decisions* is the interval on the x-axis of measured concentrations where the Agency cannot tell, with the requisite confidence, whether the actual concentration is above or below its action limit. This gray zone is better indicated by the *rectangle* between the dotted vertical lines drawn through the intersections between the power curves and the dotted horizontal lines that represent the decision error limits. This is in fact the way “gray region” is defined in EPA QA/G-4, “Guidance for the Data Quality Objectives Process.”

* To see that the area of the existing “gray zone” is uninformative, consider the situation where a perfectly accurate measurement is made once every six days for an annual-mean standard. In this case the performance curves for positive and negative bias coincide, but there is still sampling uncertainty about the actual mean value, which shows up in the curve’s non-vertical slope.

A secondary source of confusion is the numerical example offered: “for an estimate that truly is 17 ug/m^3 and the measurement system has a 10% negative bias, then 50% of the observed estimates will be declared to be less than the 15 ug/m^3 action limit.” The situation described generates data identical to those from a true value of 15.3 ug/m^3 measured with *no* bias, and something is fishy if 50% of all unbiased measurements yield estimates below 15 ug/m^3 for a true value of 15.3 ug/m^3 ! This may seem a minor discrepancy in practical terms, but it hardly contributes to a reader’s confidence that he fully understands the concept it is intended to help explain.

3. The details of the DQO modeling in Attachment 3 are sometimes murky. The values given for seasonality ratio and population variability at the bottom of page 4 don’t agree with those in the “Sel.” Column of Table 1. The first few columns of Table 1 have apparently been mislabeled in copying from Table 4-1 in Appendix 3A. It is not clear why the relatively high value chosen for the $\text{PMc}/\text{PM}_{2.5}$ ratio should be considered “conservative,” since $\text{PM}_{2.5}$ acts as an interference in the PMc measurement. And given the *a priori assumption* of zero autocorrelation, it is premature to conclude that sampling frequency is one of the factors “gray zones are most sensitive to”.

Dr. Yousheng Zeng

Comments by Yousheng Zeng

1. Among the five PMc measurement approaches, the dichotomous beta gauge method appears more attractive than the others. It is automated and provides continuous (strictly speaking, semi-continuous) measurements as opposed to manual measurements of the other two filter-based methods (FRM and the sequential dichotomous samplers). Compared to the FRM, the dichotomous beta gauge method does not require two co-located samplers. It also minimizes the potential errors and costs associated with filter handling and weighing. Compared to the two non-filter-based continuous methods, the samples of the dichotomous beta gauge method may be preserved for a further analysis of physical and chemical properties. The performance of the dichotomous beta gauge method tracks FRM significantly better than the other two continuous methods. The weakness of the dichotomous beta gauge method appears in the PM_{2.5} area – it overestimates PM_{2.5}. I would like to offer the following thoughts that may or may not help overcome the weakness.

The overestimation may be caused by (1) the different patterns of beta ray attenuation between fine and coarse particles (due to the particle size and other physical and chemical characteristics) or (2) the difference in particle mass load level between the fine and coarse channels. If reason (1) is the dominating factor, it may be difficult to overcome the overestimation issue. However, if reason (2) is the dominating factor, we may be able to improve the method by increasing the mass load. Specific implementation of this idea of increasing the mass load may be dependant on the linearity range of the beta attenuation-particle mass load curve (see Figure 1 for illustration).

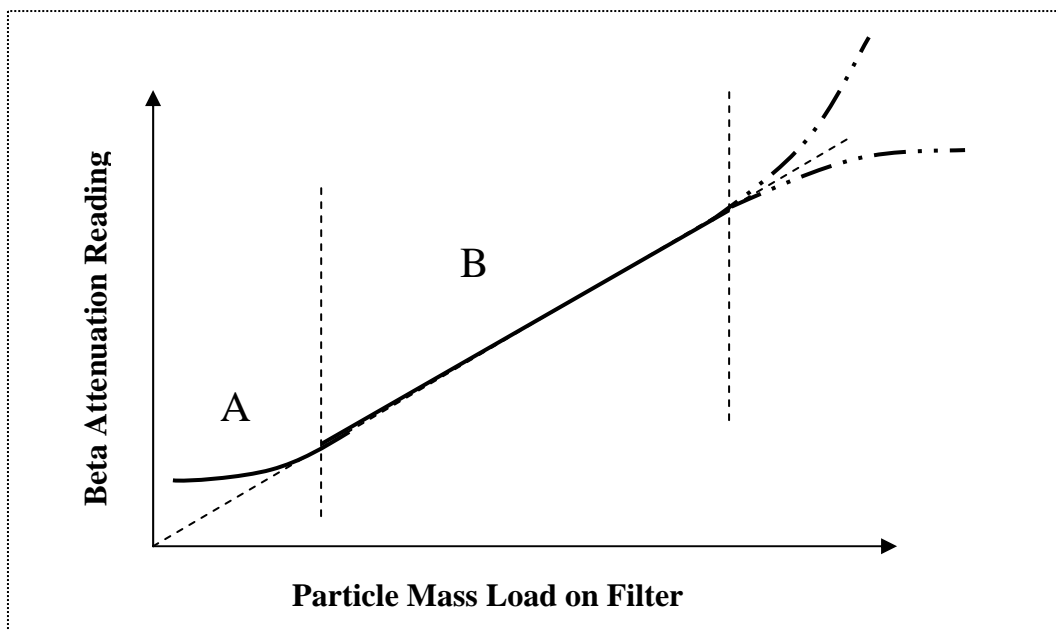


Figure 1. Hypothetic response curve of beta gauge monitor.

It is possible that the overestimation is caused by the non-linear response (overly sensitive response) in the low mass load region of the curve (Region A in Figure 1). Depending on the upward extent (towards high mass load) of the linear region of the curve (Region B in Figure 1), the above idea may be implemented differently.

If the linear range extends upward (towards high mass load) sufficiently, we may evaluate the feasibility of changing the 1-hour cycle time to 3-hour cycle time. The mass load of the PM_{2.5} filter will be tripled and hopefully enter into the linear range. A longer cycle time (still less than 24 hours) can also be evaluated. This will reduce the time resolution of the method, but should not cause any problem because the standard is on a 24-hour basis. Another possible consideration is to increase the overall sample flow without increasing the diameters of the nozzles that deposits the particles to the filters, therefore increasing the mass load to the filters.

If the linear range does not extend upward far enough or we don't want to change the upper bound, we may consider increasing the cycle time for the fine particle channel only. This should be feasible because the mechanism for advancing the filter tapes for the fine and coarse channels can be separately controlled.

Linearity is not an issue for the other two filter-based methods because they rely on gravimetric measurements. For a beta attenuation based method, linearity is an important factor. If the EPA or the manufacturer has the linearity data, it will be interesting to review the data to assess the feasibility of the above idea. If the linearity data is not readily available, the detailed data collected during the EPA PM_c multi-site evaluation may offer some clues. In this case, the "true" mass load can be derived from the co-located FRM. Hopefully the data sets are large enough and cover a reasonably wide range of mass load. If the above hypothesis is true, we should see less overestimation during days when the PM_{2.5} concentrations are high than the days when the PM_{2.5} concentrations are low.

In summary, if the overestimation of PM_{2.5} is related to the linearity of the beta attenuation response to the particle mass load, the weakness may be overcome by manipulating the amount of particle mass deposited on the filter tape using one or a combination of the above approaches.

2. The intrusion of coarse mode aerosols into the fine channel (major flow) in virtual impactors was discussed in the study. The issue is applicable to the sequential dichotomous samplers, the dichotomous beta gauge samplers, or the TEOM samplers. Has there been any study to evaluate whether or not the intrusion can be minimized by increasing the diameter of the entrance of the coarse channel (minor flow) while maintaining the same major/minor flow split? The separation of PM_{2.5} and PM_c is determined by the geometry of the separator and the face velocity at the entrance of the coarse channel. The minor flow is just a carrier to transport the already separated coarse particles to the coarse filter and its volumetric flow rate should not be critical for coarse particle separation. The opening of the coarse channel (minor flow) can be larger than the nozzle above it. A larger opening may prevent coarse particles from slipping into the

fine channel. It should not increase the amount of fine particles entering the coarse channel as long as the volumetric flow rate into the coarse channel (minor flow) remains the same because the fine particles should be evenly distributed within the air flow. In the extreme case where the minor flow rate is reduced to zero, the coarse particles should still be captured in the coarse channel (essentially same as the 2.5 μ impactor well in the sampler of FRM for PM_{2.5}). This analysis also leads to another related topic. If this analysis is valid, the split between the major flow and the minor flow should be a little more flexible (as opposed to 9:1) as long as both flows can be measured to the degree of desired accuracy to keep track of partition of the fine particles between the two filters for correction to the filter weights.

3. The Tisch SPM-613D dichotomous beta gauge samplers use mass flow sensors to control the flow in the major and minor channels. In Section 2.3 of the PMc multi-site evaluation report, it is stated that "... however, the effect of this lack of volumetric flow control is minimal if ambient conditions do not differ substantially from those existing during the flow calibration". Is data available to support this statement? How will the diurnal temperature changes affect the accuracy of the desired volumetric flows? For a temperature change from 70 °F to 100 °F during a day, the volumetric flow can be impacted by about 6%. The face velocity at the separation point will also change by about 6%. Depending on how steep the separation curve is, this kind of change should be evaluated quantitatively. Is there any reason that this sampler design cannot use the volumetric flow control?
4. I like the EPA DQO Tool and consider it very useful. I have some questions or comments on the data input to the PMc DQO Tool. Some input data was derived from the PM₁₀ and PM_{2.5} data in the EPA Air Quality System (AQS) database. In the Technical Report on Estimating Parameters for the PMc DQO Tool, data was derived from 622 sites across the nation. My question is – Do all of these sites use the modified low volume PM_{2.5} samplers for PM₁₀ (i.e., the PM₁₀ samplers used in the EPA multi-site evaluation, which is the PM_{2.5} sampler without WINS fractionator)? Current FRM for PM₁₀ (40 CFR 50, App M) does not specify a specific type of sampler. A site may continue to use a traditional high volume PM₁₀ sampler for PM₁₀ and a low volume PM_{2.5} sampler for PM_{2.5}. If significant portion of the input data to the PMc DQO Tool involves high volume samplers, and portion of the input data is derived from the multi-site study (which did not include the high volume sampler), is there a mismatch that may cause a problem? The aerodynamic characteristic (the shape of the separation curves), the performance (bias, precision, etc.), and other parameters between high volume PM₁₀ samplers and the low volume PM₁₀ samplers modified for this study are expected to be different. What level of uncertainty will be introduced by these differences?

Furthermore, if we want to fully utilize the historical data collected by traditional high volume PM₁₀ samplers for assessment of PMc, it will be valuable to evaluate the performance of the high volume PM₁₀ samplers in the same context and the relationship between the high volume PM₁₀ sampler and the low volume PM₁₀ sampler. We know that PM₁₀ samplers just mean their "cut diameter" is 10 μ and they do not separate the particles at exactly 10 μ . The actual sizes of particles allowed to enter the samplers are

determined by the “S” shape performance curve. The steeper the curve is, the better separation the samplers will have. The curves for the high volume PM10 samplers and the PM10 samplers modified from PM2.5 samplers are expected to be different. It would be interesting to review these curves during the discussion to assess the impact and usability of data. This kind of further analysis can also help bridge the gap between the historical data and future data.

U.S. EPA/SAB/CASAC/AAMM Subcommittee
Yousheng Zeng

Additional Comments Post July 22, 2004 Meeting

July 31, 2004

The following additional comments are provided:

1. As stated in my pre-meeting written comments, the DQO Tool software seems to be a very promising tool. However, if I understand the DQO Tool correctly, based on the review materials distributed for the July 22nd AAMM Subcommittee consultation meeting, one important element may be missing and improvements in this area should be feasible.

PM sampler bias is an important input parameter to the DQO Tool. The bias may be estimated based on field study data. However, if there is a known underlying physical or chemical process that may produce a significant bias, this physical or chemical process should be considered in the bias estimation. In this case, the interplay between the sampler’s fluid dynamic characteristics and the particle size distribution of the measurement target may potentially be a significant source of bias. I have created a simple spreadsheet based computer simulation program (hereafter referred to as “PM Measurement Simulator” or “PMM Simulator” and submitted as a separate file) to evaluate this type of bias. Attachment 1 to this comment includes an example of the output of the PMM Simulator to illustrate both the importance of this type of bias and the possibility of incorporating it into the DQO Tool.

The FRM and other two candidate methods for PM_c measurement are included in Attachment 1. The two candidate methods can be any methods that separate PM_{2.5} and PM_c. The separation performance curves are presented in Figure 1 of Attachment 1. In order to evaluate the bias, three hypothetical ambient air PM cases are also created and presented in Figure 2 of Attachment 1. With the hypothetical samplers and ambient air cases, the theoretic bias (before considering other biases such as evaporation losses etc.) of the two candidate samplers with respect to the FRM can be evaluated and the results are presented in Table 1 of Attachment 1.

The results in Attachment 1 offer some insight and understanding of the potential sources of bias. In the case of City 1, for example, Candidate Method 1 performs well. It has less than 2% of bias from the FRM for all three particle size groups. However, when the same sampler is used in City 2, it has 13.7% bias. This example illustrates the importance of the interplay between the sampler separation curves (Figure 1) and the size distribution of the measurement target (Figure 2). It also illustrates the danger of broadly applying the bias data derived from limited field studies to all sites. A comparison between Candidate Methods 1 and 2 in Attachment 1 also quantitatively shows a general knowledge – the steeper the separation curve is, the better the separation will be. Since the FRM has a good steep curve, the candidate methods should have similar curves in order to minimize the bias.

2. As discussed above and illustrated by Attachment 1 (or the PMM Simulator), the relationship between the sampler separation curves and the size distribution of the measurement target is very important (for easy reference, we may call it “theoretic PM sampler bias”). Particle size distribution needs to be established in order to estimate the bias. However, we often don’t know the size distribution of the measurement target. This problem will significantly reduce our confidence in PMc measurement. An alternative solution to particle size distribution is to use the separation curves of the FRM as the reference and require all candidate methods to closely match the FRM separation curves. Candidate methods can be automated and can provide better time resolution. If they also match both the PM10 and PM2.5 separation curves of the FRM, we will have much more confidence in these methods. A simulation using the PMM Simulator will show that the theoretic PM sampler bias will be eliminated no matter what the ambient PM conditions are if a candidate method has the same separation curves as the FRM. Because there are no true values in PM measurements, the FRM measured value is considered true value. Therefore the bias issue unique to the PM measurement can be reduced to and managed as the same general bias issue in other air pollutant measurements such as SO₂ monitoring.

In practice, separation curves of a candidate method probably will not exactly match the separation curves of the FRM. EPA may consider compiling a set of ambient PM particle distribution curves that represent a wide range of possibilities. With the compilation of such size distribution data and the PMM Simulator, the potential bias of the candidate method can be evaluated and a tolerance level for deviation of the candidate separation curves from that of the FRM can be assessed. This can be accomplished by building a size distribution database in the DQO Tool and incorporating the PMM Simulator into the DQO Tool.

3. The approach discussed above and the PMM Simulator may also help us better understand vast historical PM monitoring data collected using high volume PM sampler. In order to do this, separation curves for these samplers will be needed (I assume that the data may be obtained from the manufacturers). If this approach is incorporated into the DQO Tool, the DQO Tool will be more comprehensive. Also, EPA may consider adding a few typical high volume PM10 samplers in the next field (or lab) study for the purpose

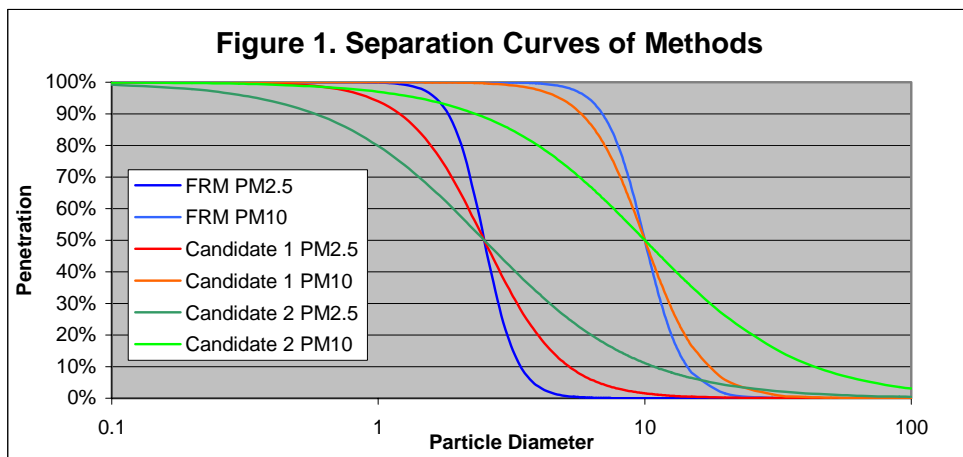
of possibly bridging the gap between the historical PM data collected by high volume samplers and the data collected by new samplers.

4. As a follow-up to my pre-meeting written comment No. 2, I made rough calculations based on the nozzle dimension data provided by Mr. Tom Merrifield during the July 22nd meeting. The acceleration nozzle I.D. (D1) of the Kimoto impactor is 3.2 mm, the coarse receiver nozzle I.D. (D2) is 4.0 mm, and the nozzle to nozzle distance (S) is 4.0 mm. The ratios of D1/D2 and D1/S are 0.8, relative close to 1. The similar ratios in the PM2.5 impactor in the FRM are much smaller (D1/D2 ~0.14; D1/S ~0.3; these numbers are not confirmed, but the differences are larger enough to make a qualitative statement). I just want to reiterate my pre-meeting comment No. 2. Increasing D2 may reduce coarse particle intrusion and make the separation curve steeper (see my pre-meeting written comment No. 2 for rationale).

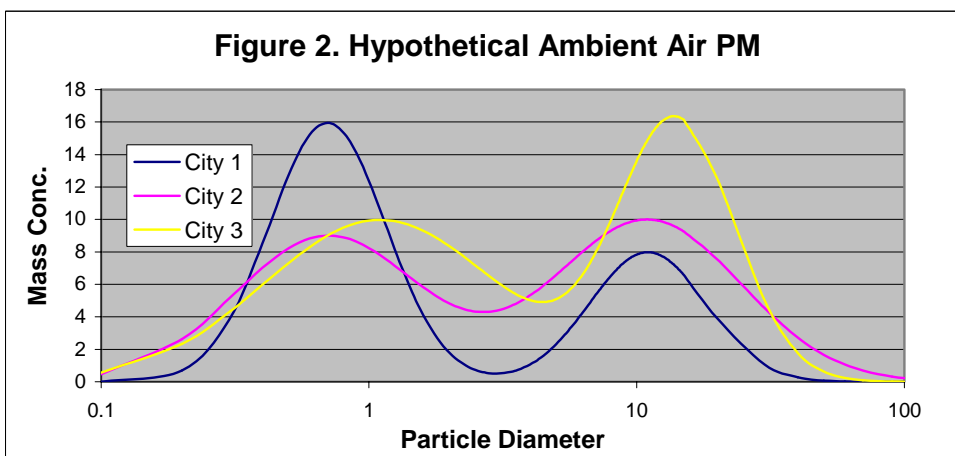
Attachment 1. Relationship Between PM Sampler Separation Curves And PM Monitoring Results

Assumed Particle Separation Curves of PMc Measurement Methods

Each PMc measurement method has two separation curves, one for PM10 (i.e., sampler inlet) and one for PM2.5 (i.e., WINS in FRM or virtual impactor in Dichotomous type sampler). For illustration purpose, the following separation curves are used. Although the actual curves may be different; the conclusions drawn from this illustration should be valid.



Three Hypothetical Cases for Ambient Air PM



Simulated Monitoring Results

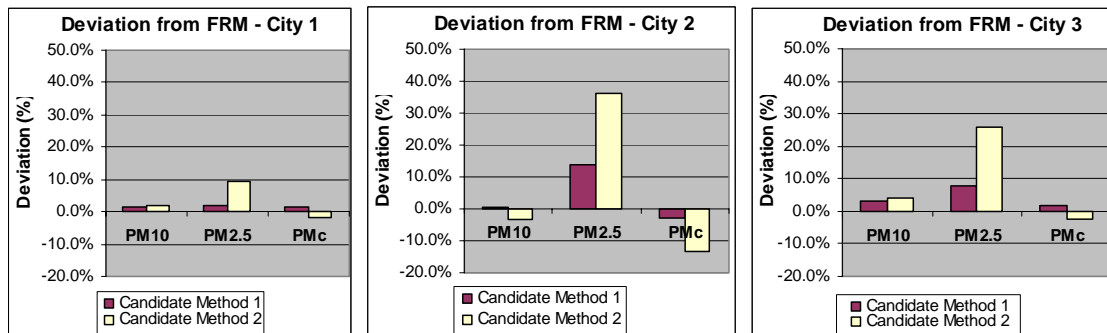
If the above PMc measurement methods are used in the three cities, the following results are expected:

Table 1. Simulated Monitoring Results

City/PM	FRM	Candidate Method 1			Candidate Method 2		
		Result	Diff fr FRM	% Diff	Result	Diff fr FRM	% Diff
City 1							
PM10	49.0	49.8	0.8	1.6%	49.9	0.8	1.7%
PM2.5	15.5	15.7	0.3	1.8%	16.9	1.5	9.5%
PMc	33.6	34.1	0.5	1.5%	32.9	-0.6	-1.9%
City 2							
PM10	74.1	74.5	0.3	0.5%	71.7	-2.4	-3.3%
PM2.5	15.2	17.3	2.1	13.7%	20.7	5.5	36.0%
PMc	58.9	57.2	-1.7	-3.0%	51.0	-7.9	-13.4%
City 3							
PM10	85.9	88.5	2.6	3.0%	89.4	3.5	4.1%
PM2.5	19.6	21.1	1.5	7.7%	24.7	5.1	25.7%
PMc	66.3	67.3	1.1	1.6%	64.8	-1.5	-2.3%

The results are also highlighted in Figure 3.

Figure 3. Deviations (%) of the two candidate methods from FRM in three hypothetical cities



Particulate Matter Measurement Simulator (PMM Simulator)

Created by Yousheng Zeng, Member of Ambient Air Monitoring and Methods (AAMM) Subcommittee
U.S. EPA Science Advisory Board (SAB), Clean Air Scientific Advisory Committee (CASAC)

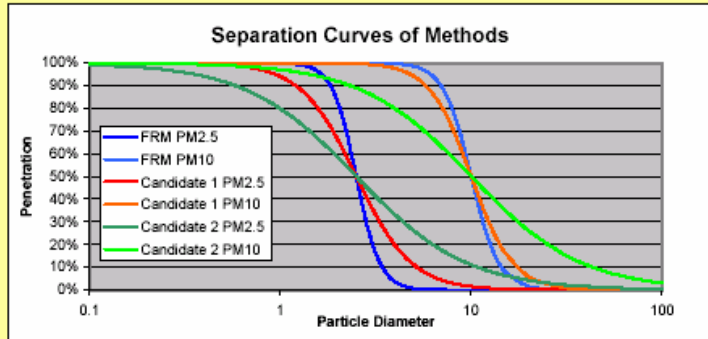
Preliminary - Verification or refinement (e.g. correction for fines in coarse channel) may be needed
July 2004

Input

Measurement Method		PM2.5	PM10
Federal Ref Method (FRM)	D_{cut}	2.5	10
	P	7	6
Candidate Method 1	D_{cut}	2.5	10
	P	3	4
Candidate Method 2	D_{cut}	2.5	10
	P	1.5	1.5

D_{cut} = Cut diameter of PM separation mechanism of the method.

P = Power parameter to determine how steep the separation curve is.



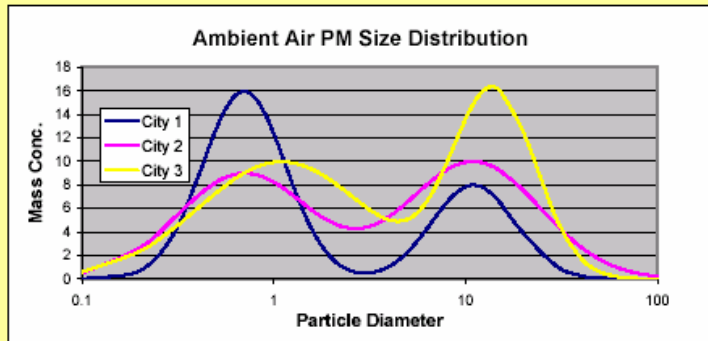
Ambient Air PM Size Distribution		D_{peak}	Sigma	A
City 1	Fine mode	0.7	0.5	20
	Coarse mode	11	0.5	10
City 2	Fine mode	0.7	0.8	18
	Coarse mode	11	0.8	20
City 3	Fine mode	1.1	1	25
	Coarse mode	14	0.5	20

D_{peak} = Particle diameter corresponding to the peak of the particle size distribution curve for each mode.

Sigma = A parameter describing how wide the particle size distribution curve spreads.

A = Amplitude parameter to determine the height of each particle size distribution curve (simulated PM concentration level).

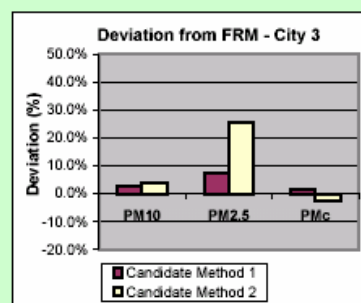
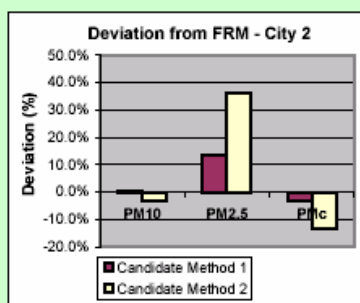
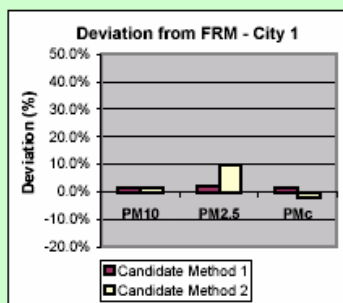
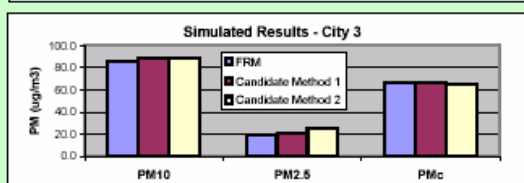
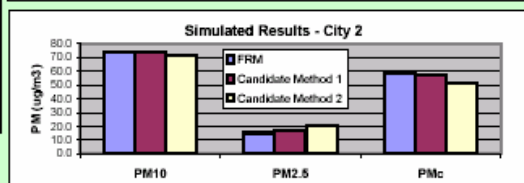
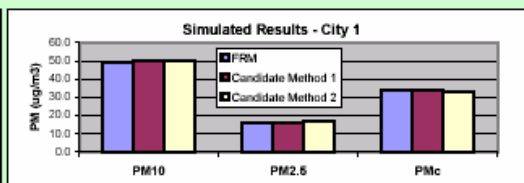
Caution - Because the distribution curves are on log scale, the magnitude of the fine mode is smaller than it appears.



Simulated Monitoring Results

If the above PMc measurement methods are used in the three cities, the following results are expected:

City/PM	FRM	Candidate Method 1			Candidate Method 2		
		Result	Diff fr FRM	% Diff	Result	Diff fr FRM	% Diff
City 1							
PM10	49.0	49.8	0.8	1.6%	49.9	0.8	1.7%
PM2.5	15.5	15.7	0.3	1.8%	16.9	1.5	9.5%
PMc	33.6	34.1	0.5	1.5%	32.9	-0.6	-1.9%
City 2							
PM10	74.1	74.5	0.3	0.5%	71.7	-2.4	-3.3%
PM2.5	15.2	17.3	2.1	13.7%	20.7	5.5	36.0%
PMc	58.9	57.2	-1.7	-3.0%	51.0	-7.9	-13.4%
City 3							
PM10	85.9	88.5	2.6	3.0%	89.4	3.5	4.1%
PM2.5	19.6	21.1	1.5	7.7%	24.7	5.1	25.7%
PMc	66.3	67.3	1.1	1.6%	64.8	-1.5	-2.3%



NOTICE

This report has been written as part of the activities of the Environmental Protection Agency's (EPA) Clean Air Scientific Advisory Committee (CASAC), a Federal advisory committee administratively located under the EPA Science Advisory Board Staff that is chartered to provide extramural scientific information and advice to the Administrator and other officials of the EPA. The CASAC is structured to provide balanced, expert assessment of scientific matters related to issues and problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the EPA, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use. CASAC reports are posted on the SAB Web site at: <http://www.epa.gov/sab>.