

Interactive comment on “Application of absolute principal component analysis to size distribution data: identification of particle origins” by T. W. Chan and M. Mozurkewich

T. W. Chan and M. Mozurkewich

Received and published: 23 December 2006

We do not agree with the referee’s recommendation that we combine the two papers. This paper is not “very short”; it will be about 5 pages in print. If we were to combine the two MS, either the resulting paper would be very long and unwieldy or we would be forced to cut out a great deal of detail, making the result very difficult to understand. Since we are not simply applying some established method in analyzing our data, we need to provide enough technical detail for other researchers to be able to apply the method. The two papers also have very different objectives. The first deals with using weighted PCA to create a greatly simplified representation of a size distribution data set; that analysis does enable source identification but is a necessary prerequisite to the analysis in the second paper. The analysis in this paper uses the simplified

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

representation in an unweighted, scaled PCA to identify sources.

We are puzzled by the referee's complaint that the paper "discusses localized results". It seems to us that this would apply to virtually field studies. There is also a broader purpose here, but the referee seems to have missed it since he writes: "Summarizing what I miss is the generic aspect of the MS." In the abstract we state that "Here we demonstrate how to combine the simplified size distribution data with trace gas measurements and meteorological data to determine the origins of the measured particulate matter using absolute principal component analysis"; and in the conclusions "The consistent results among different field measurements show that when combining particle number concentrations with different trace gas measurements and meteorological data, absolute principal component analysis can be useful in providing physical meaningful factors for interpretation" and "In summary, this study shows that principal component analysis can be effectively applied to data sets including size distribution data to provide useful information on the sources and origins of measured particulate matter." The technique described in this MS is certainly generic and should be applicable to a wide range of data. We demonstrated this by applying the method to four independent field data sets that were taken at different locations and different times.

The referee complains that "the method used for data evaluation is not explained nor summarized in this paper and the reader and reviewer are referred to another paper that gives all the details... I would need a 'popular' summary." We are not sure what the referee wants. We could include a few sentences summarizing the logic behind PCA.

The referee complains that "the present application paper does not make it clear what fractions were used ... in the cited references there was a very visible choice ... because these are connected to a scientific or environmental issue ... for instance UF and PM_{2.5} were used with the aim to factorise these with other well-known fractions aerosol parameters like blackness." We are not sure that we understand the referee; it sounds as if he wants us to first decide what we are looking for and then preselect only the data that would support that. We prefer to let the data speak for itself, to the extent

that is possible. Accordingly, as stated clearly in the paper, we use ALL the information in the size distribution data. The present analysis makes that possible, unlike previous attempts to apply PCA to size distribution data.

The referee complains that "In this MS the naming of the various fractions appears to be based on the local circumstances and it is not explained if such an approach is a generic one in the sense that it can be extrapolated generalized to/for other locations." The methodology should be generic but, clearly, any local data set will have its own unique features along with more generic ones. That is just what we see here: some components appear at all the sites and some are local.

Interactive comment on Atmos. Chem. Phys. Discuss., 6, 10493, 2006.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper