Atmos. Chem. Phys. Discuss., 6, S5119–S5126, 2006
www.atmos-chem-phys-discuss.net/6/S5119/2006/

Atmospheric Chemistry and Physics Discussions

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

Anonymous Referee #1

Received and published: 6 December 2006

Review of Chan and Mozurkewich, 'Application of Absolute PrincipalComponent Analysis to Size Distribution Data: Identification of Particle Origins'

This paper applies a technique described in a companion paper (ACPD 6: 10463-10492, 2006) to four field studies. The technique is used to simplify the representation of highly time and size-resolved size distribution data, by finding a small set of components that capture most of the variance of the dataset. The technique is expanded here to also analyze gas-phase species and meteorological data, jointly with the output of the size distribution analysis. The paper is interesting and appropriate for ACP, and I recommend it for publication once the issues below have been addressed.

Comment on Previous Review

The first review of this paper (posted online before my review) concludes that the present paper is not different enough from the companion paper, and suggests that the current paper should be included as an appendix of the companion paper. Sometimes authors break up their work into too many papers in order to have the same amount of work count as a larger number of publications. However I don't think this is the case here, and I strongly disagree with the recommendation from the first reviewer. The current paper applies the technique (developed in the 1st paper) to several real-world datasets, and also expands it by including gas-phase and met data. The current paper is interesting and goes significantly beyond the 1st paper. If this material was included as an appendix or continuation of the other paper, it would make that 1st paper overly long and hard to read. It would also force this paper to be condensed, when what is needed in my opinion is to add some more information and link it better to the literature. I thus support publication of both manuscripts as independent ACP papers, provided that the points below are addressed.

Main Points

P10496: the differences and appropriateness of the assumptions (to atmospheric analysis) of PCA vs. PMF-type methods should be discussed in more detail. See our review of the companion paper for details.

P10501 and 10516: Figure 2 shows that wind speed has been apportioned by the model as if it was another tracer. I don't think that this is appropriate. The various aerosol and gas-phase tracers are all 'material' and obey conservation of mass as they interconvert. Wind speed will have some correlation with some tracers, but apportioning it to the components seems a philosophical stretch to me. This issue should be at least discussed in the paper. The developer of PMF has addressed this issue by creating the 'multilinear engine' (ME) model (Paatero, 1999), in which wind speed and similar meteorological tracers are allowed to INFLUENCE the apportionment, but are

ACPD
6, S5119–S5126, 2006

EGU

NOT apportioned themselves. That seems more appropriate to me. Is there a significant difference in the results when wind speed is included or excluded from the APCA analysis?

P10502/L14: The discussion of new particle formation and growth events is incomplete and not completely aware of the recent literature:

- It is not surprising that SO2 itself is not correlated with the nucleation events. As long as there is some SO2, the relevant species is sulfuric acid (SA). Since the gas-phase oxidation of SO2 is slow (time scale of days), the concentration of SA doesn't depend very strongly on the amount of SO2, but it depends more strongly on the concentration of the OH radical and of aerosol surface area. Peter McMurry (2005) has shown that it is the tradeoff of those two factors (condensable, often SA production, and condensational sink) that typically controls nucleation. The anticorrelation of nucleation with the concentrations of the larger particle sizes (where the surface area is larger), apparent especially in Fig. 2c, is an indication of the later effect and is also consistent with the cited paper by Kulmala. This should be discussed in more detail.

- Similarly Stanier et al. (2004) studied the climatology of new particle formation in Pittsburgh. Pittsburgh is only 200 hundred miles to the south of the sites studied here, and their findings are likely relevant to the locations discussed in this paper. They also showed that SO2 * UV light can be a useful surrogate of the H2SO4 production rate. This could be evaluated here, rather than using SO2 and radiation separately.

- Zhang et al. (2004) showed that the composition of new particles formed in Pittsburgh were dominated by sulphuric acid. All condensable species available (NH3, SOA, nitrate) later condensed on the growing new particles. It is very likely that the dynamics of new particles formation at these Canadian sites are similar.

- There are several other published studies at these and nearby locations (e.g. Ruphaketi et al. (2005), Broekhuizen et al. (2006), etc.) , and the present results should be placed in the context provided by those prior studies.

ACPD

6, S5119–S5126, 2006

P10501-10506, section 4: I do agree with the previous reviewer that this description is at times hard to follow as it jumps between sites and components. This is partially unavoidable due to the 'matricial' nature of the results. I invite the authors to think of ways to clarifying this presentation, including perhaps a 'matricial' table of the results along the lines above, or perhaps another map with similar information (but with most geographic features such as roads being removed).

P10501, section 4: the previous paper (P10480/L27) states that 'mixed' components represent atmospheric processing. This paper discusses atmospheric processing but does not identify which components are mixed. It would be useful if this was done, and if the eigenvalues associated with each component were reported.

P10501, section 4: the average size distribution of each component should be reported. This would allow the reader to see that the processes invoked here are consistent with the size distributions observed in other studies.

P10517/Fig3: this figure shows that the photochemical component scores are high when it is cloudy and tend to be flat when it is clear. This seems backwards with the interpretation given here, and should be addressed.

Detailed Points

General question: Is it possible to put physical units on loadings and scores? Or are the units arbitrary? Either way, this should be briefly discussed.

General question: How do the scores of the factors strongly associated with particle sizes in the PCA compare to the scores of the particle sizes in the APCA?

P10494/L13-14: a component is described as 'trace-gas variations associated with boundary layer dynamics.' Presumably the particle size distributions also show significant changes when the boundary layer rises?

P10495/L1-2: a recent innovative application of receptor models to time series of aerosol mass spectra has been reported by Zhang et al. (2005), and could also be

ACPD

6, S5119–S5126, 2006

cited here.

P10496/L20: (same comment as in review of 1st paper) This would be better served with an example. How can positive/negative be arbitrary here, but have meaning about correlation when examining results? Is this changing of sign different than a rotation or linear transformation?

P10497/L11-19: have the authors tried to conduct the whole analysis (size distributions plus gases and met) in one step? If properly weighed, the results should be similar to the current two-step procedure, and could have advantages in that the detailed size dist data is allowed to interact with the other tracers.

P10499/L20: The previous paper reported 5-8 factors. Hamilton 2000 data had 5 factors with different mode diameters. Please explain this difference. Why weren't Pacific 2001 data included in this paper?

P10499/L24: the 'mixed' components are defined in the 1st paper, but are not described here. A short explanation and a reference to the other paper for details are needed.

P10500/L16: the type of 'unsatisfactory results' obtained in this case should be described in more detail (1-2 sentences). This information may be useful to other practitioners.

P10501/L2: Modified scree plots could be included in supplementary information.

P10501/L8: please explain what is meant by 'reasonable physical interpretation'

P10501/L24: There should be some discussion about the interpretation of negative values in scores, e.g. in figures 5 (no negative values) vs. 3 and 8 (with negative values).

P10503/L15: Can the variation in regional SO2 compared to variation of local SO2 be supported/explained with data?

ACPD
6, S5119–S5126, 2006

EGU

P10504/L5: there is often a difference between local wind direction and back trajectories, e.g. due to the variation of wind direction and speed with height. Have the authors evaluated whether this difference is significant here?

P10505/L6: The boundary layer dynamics factor was not found at Egbert and Ox data is not available for that site. Nonetheless, boundary layer dynamics should occur at this site. Can you offer an explanation about how this factor might be combined into the other Egbert factors?

P10505/L22: The justification for the meaning of this factor seems weak.

P10506/L18: Can some reference be given for the claim of the origin/properties of transported particles?

P10513/Table2: Could there be differences between the Hamilton and other datasets because of the difference in size ranges measured? (e.g. in the number of components needed to represent the data). Also the Simcoe and Egbert sites used a lower nominal sizing resolution (5 instead of 10). The use of lower resolution smoothes some finer features of the size distribution, and could suppress some small components in the analysis. This should be mentioned. Have the authors tried to numerically smooth the size distributions recorded at higher resolution before doing the APCA, to characterize this effect?

P10521/Fig7: Does the figure represent one or two years of data? It might be interesting to see the years in different colors.

P10521/Fig7: If the dots represent relative magnitudes, shouldn't there be one point on the probability=1 line?

P10522/Fig8: Can this be combined or overlaid with an image plot of the number concentration to show one mode grow into another? This may work better if a subset of the time series (e.g. Jul 5-9) is chosen for the X-axis.

Grammar etc.

ACPD

6, S5119–S5126, 2006

P10494/L18: ‘roles’ should be ‘role’

P10498/L11: ‘source’ should be ‘sources’

P10504/L15: ‘k’ should be ‘>=‘

P10504/L18: It would be clearer to note in line 8 that the score value of 2.5 is chosen as the predetermined threshold criterion.

References

Paatero, P., 1999. The multilinear engine–a table-driven least squares program for solving multilinear problems, including the n-way parallel factor analysis model. Journal of Computational and Graphical Statistics 8, 854-888.

Zhang et al. Deconvolution and quantification of hydrocarbon-like and oxygenated organic aerosols based on aerosol mass spectrometry. Environmental Science and Technology, 39: 4938-4952, 2005

McMurry et al. A criterion for new particle formation in the sulfur-rich Atlanta atmosphere. JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES 110 (D22): Art. No. D22S02 NOV 4 2005

Stanier et al.. Nucleation events during the Pittsburgh air quality study: Description and relation to key meteorological, gas phase, and aerosol parameters. AEROSOL SCIENCE AND TECHNOLOGY 38: 253-264 Suppl. 1, 2004

Zhang et al. Insights into the chemistry of new particle formation and growth events in Pittsburgh based on aerosol mass spectrometry. ENVIRONMENTAL SCIENCE TECHNOLOGY 38 (18): 4797-4809 SEP 15 2004

Rupakheti et al. An intensive study of the size and composition of submicron atmospheric aerosols at a rural site in Ontario, Canada. AEROSOL SCIENCE AND TECHNOLOGY 39 (8): 722-736 AUG 2005

ACPD

6, S5119–S5126, 2006

Closure between measured and modeled cloud condensation nuclei (CCN) using size-resolved aerosol compositions in downtown Toronto. Broekhuizen, K. , Chang, R.Y.-, Leaitch, W. R., Li, S.-M. and Abbatt, J. P. Atmospheric Chemistry and Physics, Vol. 6, pp 2513-2524, 30-6-2006.

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