Atmos. Chem. Phys. Discuss., 8, S6424–S6432, 2008
www.atmos-chem-phys-discuss.net/8/S6424/2008/
© Author(s) 2008. This work is distributed under
the Creative Commons Attribute 3.0 License.

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Interpretation of organic components from positive matrix factorization of aerosol mass spectrometric data" *by* I. M. Ulbrich et al.

I. M. Ulbrich et al.

Received and published: 27 August 2008

Response to Anonymous Referee #1

General Response

We thank Anonymous Referee #1 for his or her comments on our work. The referee says that this paper "is an important contribution to the field," which implies that he/she supports its publication in ACP. Most of the comments are very general and are not stated as in opposition to most of the current manuscript. We have repeated the comments here in italics and added comment numbers for easy reference between points in the response. Our replies follow each excerpt. Changes to the manuscript text are presented in bold.

Detailed Response to Individual Comments

1. This paper is an important contribution to the field. Positive matrix factorization for Aerodyne aerosol mass spectrometry data has so far only been used by Lanz et al. (2007, see reference in paper). However one can expect a much extended use of PMF by this aerosol mass spectrometer community, especially when an easy applicable code like the one presented here becomes available. Both this paper and the paper by Lanz et al., (2007) emphasize the necessity to carefully use the method and the need for as much evaluation of the results as possible. I think that both examples should be a role model for some of the future work in source apportionment using aerosol mass spectrometer data.

[Response]: We thank the Referee for these positive comments. The Referee's comment about the extended use of PMF is quite relevant, as several additional studies including PMF analysis of AMS data have already been published (Aiken et al., ES&T, 42, 4478-4485, doi: 10.1021/es703009q, 2008; Nemitz et al., Aerosol Sci. Technol., 42, 636-657, 2008; Herndon et al., 35, L15804, GRL, doi: 10.1029/2008GL034058, 2008; Docherty et al., ES&T, in press, 2008; Cubison et al., ACP, in press, 2008) which present results of PMF of AMS data for datasets in Mexico City, Riverside, CA, and Boulder, CO. Cottrell et al. (JGR, 113, D08212, 2008) refer to consistent results obtained between multiple component analysis (MCA) of Zhang et al. (2007) and PMF analysis of an AMS dataset from Thomson Farm, NH.

Lanz et al. (2007) give a detailed explanation of the factors in the 2- to 7-factor solutions of their dataset and discuss the criteria for evaluation of solutions. Our paper further evaluates the possibilities and limits of analyzing AMS data with PMF. We present PMF analyses of a real and synthetic AMS datasets and explore the solutions as a function of number of factors and rotation and investigate the retrieveability of small factors. We believe that it is a solid new contribution that deserves to be published in the peer-reviewed literature.

ACPD

8, S6424–S6432, 2008

2. The case study of Pittsburgh is very well described and it seems very important to show that very similar factors, especially OOA1 and OOA2 correlating with sulfate and nitrate respectively as found by Lanz et al. (2007). It would be very nice to discuss how well the OOA1s and OOA2s from both studies compare. This might be done also in a future paper including data from more stations.

[Response]: We already address part of this issue in the ACPD paper (P6750 L5-7) with the following text "A similar OOA-II factor with a less oxidized spectrum and a high correlation with nitrate was reported by Lanz et al. (2007) for their dataset in Zurich in summer of 2005. These authors also interpreted OOA-II as fresh SOA."

At this point only a qualitative comparison of the spectra is possible, since the spectra from Lanz et al. have not been submitted to the AMS spectral database. Our Pittsburgh OOA-II looks more like the combined OOA-II by Cottrell et al. (2008) than the OOA-II presented by Lanz. The OOA-I's are similar, and we will change the current text (P6749, L12-14) which currently states, "The OOA that accounts for most of the mass is very similar to that identified by Zhang et al. (2005a) and is now termed OOA-I, following the nomenclature of Lanz et al. (2007)." to "**The OOA that accounts for most of the mass is very similar to that identified by Zhang et al. (2005a) and Lanz et al. (2007), and is now termed OOA-I, following the nomenclature of Lanz et al. (2007).**"

A more detailed comparison of the different OOA-Is and -IIs obtained in different studies is beyond the scope of this paper, and should be explored in a future publication. It would certainly be more interesting to compare multiple OOAs when a larger database of cases in many different environments has been published. We note that at least in some cases the OOA-x appear to be interpolants which PMF finds to represent a continuously increasing degree of oxidation/aging of the SOA. (And as we argue in the paper, (P6752 L10-14) some part of the evolution may not be representable by this type of interpolation and likely leads to the extra residual that we observe in this study.) Thus the specific OOAs found in each study also relate to the distribution of

ACPD

8, S6424–S6432, 2008

Interactive Comment

precursors and photochemical ages that happen to be sampled in a particular study. This complicates the interpretation of a direct comparison. A mention of this possible effect will be included in the revised version of the paper.

3. In general, it should be mentioned here that also with this paper, only two field campaigns (Pittsburgh and Zürich) have been assessed with PMF in detail. It should be emphasized that this should be done for more stations and campaigns in future before drawing too many general conclusions about the general applicability of the PMF method.

[Response]: As we note in the response to comment 1 above, PMF has also been applied to 6 other papers which are now published or in press. Two of these were already cited in the ACPD version, and we will add citations to the other papers in the revised version. Although, to avoid repetition, none of these papers include the level of detail about the PMF solutions reported in this work and that of Lanz et al. (2007), an equally rigorous examination of the dataset solutions was made to choose the best solution in each case. In addition to these datasets, our groups have papers in preparation for approximately 30 additional datasets analyzed by PMF. We and others are finding that PMF is often a useful method for analysis of OA components in AMS data.

We agree that we did not make clear enough in the ACPD version that the conclusions of this analysis are based on this Pittsburgh case, and that similar analyses should be carried out for other datasets to further explore the generality of the conclusions. Note that we did address this point in response to the comment from B. Resson in this public discussion, before any of the reviews were posted, with the following text: "If we wanted to expand the current paper, we think that rather than the studies suggested by B. Resson, a more relevant expansion would be to repeat the type of analyses in the paper using other urban background datasets from different cities and seasons, to further explore the generality of our conclusions."

ACPD

8, S6424–S6432, 2008

CC BY

We will add text to the abstract, discussion, and conclusions of the paper addressing this point. For example the following text will be added to the abstract: "**Researchers are urged to analyze future datasets carefully, including synthetic analyses, and to evaluate whether the conclusions made here apply to their datasets.**"

4. The part regarding the synthetic dataset seems to be less convincing so far. There are different ways how one could produce the synthetic dataset .. so it might be dangerous again to draw too many too general conclusions. I would strongly support the opinion of Paatero to split the paper in two parts. I also agree with Paatero that most figures are not easily legible. The reduction to the Pittsburgh study here and an extension to the discussions of the synthetic datasets in a second paper would aid to increase figure size and clarity in the individual papers.

[Response]: We agree that there are multiple ways to produce a general synthetic dataset. However few of them would be relevant to the study here — exploring the application of PMF to AMS organic datasets. We therefore carefully chose factors and characteristics to create the synthetic datasets presented in which:

- Factor MS and TS come from two different factor analysis methods applied to an AMS organic dataset and actual MS obtained from AMS instruments. The factors used to create synthetic datasets are therefore representative of the types of factors that have so far been retrieved from such methods (including the Lanz et al. results). The two-factor base case and all three-factor cases are created with MS factors which reflect the structure of "near-zero" values typical in AMS spectra. The TS factors have time variations which are qualitatively similar to those in other studies. However, factor TS may vary across AMS datasets due to local sources, photochemistry, meteorology, etc. This point is addressed by the text added in response to comment #3 above about the need to study additional cases.

- Values of the measurement standard deviation for the synthetic datasets are estimated by a method identical to the method used for real data. The "closed beam"

ACPD

8, S6424–S6432, 2008

spectrum is estimated from actual and average values obtained from the same instrument used during the real campaign, in order to use as much actual, characteristic AMS data as possible.

- Noise is estimated for the synthetic datasets that reflects our best understanding of the actual noise in AMS measurements, including noise created from counting statistics and electronic noise.

We believe that the above assumptions are the most reasonable and produce the most realistic synthetic data that could be generated for this case study. The reviewer does not provide any specific suggestions about the synthetic data generation process for us to address in the revision of the paper. However, we will address the specific criticisms of the synthetic datasets analysis given by Referee P. Paatero in more detail in that response.

Specifically in response to the suggestion of splitting the paper in two, we will remove from the revised version Figures 11 and 12 and the associated text in sections 3.2.1, 3.2.2, Discussion, and Conclusions. These figures and sections dealt with the retrievability of factors with different degrees of correlation. P. Paatero indicated that the number of zeros in the factors should also be taken into account, which is consistent with our previous experience. We have performed preliminary analysis along those lines but find that the complexity of this topic is too large to do it justice in the present paper. Thus, time permitting, we will explore these topics in a future publication.

We will be sure to ask ACP to print the figures in as large and clear format as possible in the final version. The small size in the ACPD version is an unfortunate effect of fitting the figures onto the landscaped, half-page format used by ACPD, independent of the number of figures in the submitted manuscript.

Specific comments: 5. - p. 6746: The method, using R for m/z>44 was introduced in Alfarra et al., (2006) not in Alfarra et al., (2007). See reference below.

ACPD

8, S6424–S6432, 2008

CC BY

[Response]: We will correct this reference.

6. - p. 6747: I would also recommend to discuss in somewhat more detail, the differences between the results using the robust and non-robust mode.

[Response]: We state on pg. 6747, lines 11-12: "We explored the [...] use of the robust mode in PMF (in which outliers in the iterative fit ($|e_{ij}/\sigma_{ij}|$>4) are not allowed to pull the fit with weight >4). Differences in the factor MS and TS were minor in all cases."

Comparison of the 2- to 4-factor solutions of the real Pittsburgh dataset modeled in the robust and non-robust modes are available at http://tinyurl.com/5csnnj. Because the factors obtained in the robust and non-robust modes are virtually identical, we did not feel that more discussion was worthwhile. These figures will be added to the supplemental information of the paper. It is possible that the robust vs. non-robust mode makes a larger difference in other cases, and we strongly suggest that other researchers also perform this comparison for their datasets.

7. - p. 6750 line 27: This is true for both OOA-I and OOA-II

[Response]: This comment refers to the description of the 4-factor solution of real data in which "we warn about trying to interpret e.g. one of the OOA-Is as "biogenic" and the other as "anthropogenic" or similar splits, in the absence of strong evidence to support these assignments."

The statement in this location is meant specifically to refer to the apparent "splitting" of the OOA-I factor. We do agree with the point made by the Referee, that there is insufficient support to attribute OOA-I or OOA-II to anthropogenic or biogenic sources. We will add the following to in the description of the 3-factor solution of the real Pittsburgh case on pg. 6750 at the end of line 4: "**No evidence is available to support the identification of OOA-I or OOA-II as "anthropogenic" or "biogenic" in origin.**"

8. - p. 6752 line 10; Also the OOA-I spectrum might vary with oxidation time. It also

ACPD

8, S6424–S6432, 2008

might be different at different stations..

[Response]: This comment refers to the correlation between the TS of the residual and OOA-II. It is definitely possible that the OOA-I spectrum might change with oxidation time in the atmosphere, although less likely than for OOA-II due to the likely larger atmospheric aging time of OOA-I. Also the excess Q-contribution does not correlate strongly with the OOA-I TS (Fig. 7). The OOA-I MS is substantially more stable with respect to rotation than the OOA-II MS (Fig. 9), which implies that the OOA-I factor is very stable.

Whether OOA-I's vary between locations is not relevant to the discussion of the residual of this specific dataset at this point in the paper. It is true that OOA-I's at different locations may not be identical, but we note that the OOA-I's calculated for Zurich and Pittsburgh have similar MS. The OOA-I reported by Nemitz et al. (2008) and Cottrell et al. (2008) are also similar. See also our response to point 2 above about the possibility that the OOAs returned by PMF are dependent on the distribution of photochemical ages sampled in a given study.

References

A.C. Aiken, P.F. DeCarlo, J.H. Kroll, D.R. Worsnop, J.A. Huffman, K. Docherty, I.M. Ulbrich, C. Mohr, J.R. Kimmel, D. Sueper, Q. Zhang, Y. Sun, A. Trimborn, M. Northway, P.J. Ziemann, M.R. Canagaratna, T.B. Onasch, R. Alfarra, A.S.H. Prevot, J. Dommen, J. Duplissy, A. Metzger, U. Baltensperger, and J.L. Jimenez. O/C and OM/OC Ratios of Primary, Secondary, and Ambient Organic Aerosols with High Resolution Time-of-Flight Aerosol Mass Spectrometry, Environmental Science and Technology, 42, 4478-4485, doi: 10.1021/es703009q, 2008.

L.D. Cottrell, R.J. Griffin, J.L. Jimenez, Q. Zhang, I. Ulbrich, L.D. Ziemba, P.J. Beckman, B.C. Sive, and R.W. Talbot. Submicron particles at Thompson Farm during ICARTT measured using aerosol mass spectrometry, Journal of Geophysical Research-Atmospheres, 113, D08212, 2008; doi:10.1029/2007JD009192.

ACPD

8, S6424–S6432, 2008

M.J. Cubison, B. Ervens, G. Feingold, K. Docherty, I. Ulbrich, L. Shields, K. Prather, S. Hering and J.L. Jimenez. The Influence of Chemical Composition and Mixing State of Los Angeles Urban Aerosol on CCN number and Cloud Properties. Atmospheric Chemistry and Physics, in press, Aug. 2008.

K.S. Docherty, E.A. Stone, I.M. Ulbrich, P.F. DeCarlo, D.C. Snyder, J.J. Schauer, R.E. Peltier, R.J. Weber, S.M. Murphy, J.H. Seinfeld, D.J. Eatough, and J.L. Jimenez. Apportionment of Primary and Secondary Organic Aerosols in Southern California during the 2005 Study of Organic Aerosols in Riverside (SOAR). Environmental Science and Technology, in press, Jul. 2008.

S.C. Herndon, T.B. Onasch, E.C. Wood, J.H. Kroll, M.R. Canagaratna, J.T. Jayne, M.A. Zavala, W.B. Knighton, C. Mazzoleni, M.K. Dubey, I.M. Ulbrich, J.L. Jimenez, R. Seila, J.A. de Gouw, B. de Foy, J. Fast , L.T. Molina, C.E. Kolb and D.R. Worsnop. The Correlation of Secondary Organic Aerosol with Odd Oxygen in a Megacity Outflow. Geophysical Research Letters, 35, L15804, doi: 10.1029/2008GL034058, 2008.

E. Nemitz, J.L. Jimenez, J.A. Huffman, M.R. Canagaratna, D.R. Worsnop and A.B. Guenther. An eddy-covariance system for the measurement of surface / atmosphere exchange fluxes of submicron aerosol chemical species - first application above an urban area, Aerosol Science and Technology, 42, 636-657, March 2008.

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 6729, 2008.