Atmos. Chem. Phys. Discuss., 8, S2457–S2459, 2008
www.atmos-chem-phys-discuss.net/8/S2457/2008/

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: 9 May 2008

We would like to thank Prof. Paatero for his detailed comments on our manuscript (see preliminary response to main comments below). A point-by-point response will be posted later once we have time to fully address all the details. Submission of additional comments from Prof. Paatero during the public discussion phase of the review, which continues for 4 more weeks, would be greatly appreciated. We believe that the community will benefit from a timely public discussion, rather than leaving elements of the review to a 2nd round of review, which is not public.

========== Preliminary Response to Main Comments ==========

It appears to us, based on preliminary tests, that the larger issues raised by Prof.

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

Paatero can be addressed in a timely manner. Note that we can run and analyze many additional numerical experiments quickly because of our investment in the PMF Evaluation Panel (Fig. 2 in the paper). The four most important issues raised by the reviewer are: (a) The use of Pearson's R instead of uncentered R; (b) the effect of zeros on the rotatability of the solutions; (c) a mistake in the estimation of the errors; and (d) the conclusions regarding FPEAK based on the synthetic data. We can address each of these briefly here:

(a) The uncentered correlation metric that he suggests is very similar to the Pearson's R for mass spectra, and quite correlated with Pearson's R for time series (when computed with a large number of different MS and TS; see figures available at http://tinyurl.com/4cohxf). It is not yet clear to us whether changing from Pearson's R to the uncentered correlation is a worthwhile change to the manuscript. We do use an alternative metric of correlation (Pearson's R for it m/z > 44 only) on Fig.6, and as stated in the paper (P6747 L2-9) we evaluated several other metrics of correlation (Spearman's R, and custom variations of Spearman's R with e.g. threshold to eliminate very small values) and found that they did not provide significantly more information than R and $R_{m/z>44}$. There is also a benefit of consistency for staying with Pearson's R and $R_{m/z>44}$ since several previous works (e.g. Zhang et al., 2005ab; Lanz et al., 2007, 2008) already use these metrics. If we were to choose to modify the manuscript to use the uncentered correlation, this may be relatively straightforward since as described above this metric is very similar to Pearson's R for the MS and quite correlated for TS.

(b) Regarding the effect of zeros on the rotatability of the solutions, this is a topic that we have investigated, but that we decided to keep out of the present manuscript due to concerns about excessive length. As stated by Prof. Paatero, we find that when the input factors have more zeros, the solutions are less rotatable. We can include this analysis in the revised paper, and this should not cause a great delay.

(c) As pointed out by Prof. Paatero, there is a mistake in the specification of the error matrix in the "Pittsburgh Real" case (but not on the synthetic cases) due to the ap-

ACPD

8, S2457–S2459, 2008

plication of 3-point boxcar smoothing to the data to reduce high-frequency noise. We had actually realized this before this review was posted, and have already evaluated the effect of using the correct error matrix. Our evaluation of this effect indicates that all changes for the “Pittsburgh Real” case are very minor, and our conclusions can be evaluated with the figures presented in the paper. New versions of all affected figures are available from http://tinyurl.com/4klulg. There are many accompaning plots with scatter plots between the versions from the submitted paper and the updated figures available from http://tinyurl.com/4emw7y.

(d) Based on preliminary tests, it appears that even solutions generated with FPEAK $<$ or $>$ 0 still reproduce the input solution very well at FPEAK close to 0. We have already run some cases and can run additional cases examining this point, so that we can address this issue thoroughly in the revised version in a timely manner. Some preliminary figures describing these results are available from http://tinyurl.com/4uzbsl.

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