Atmos. Chem. Phys. Discuss., 8, S1607–S1609, 2008
www.atmos-chem-phys-discuss.net/8/S1607/2008/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Assessing positive matrix factorization model fit: a new method toestimate uncertainty and bias in factor contributions at the daily timescale" *by* J. G. Hemann et al.

Anonymous Referee #1

Received and published: 13 April 2008

Presented in this manuscript is a new method capable to assess uncertainty and bias in PMF source apportionment results. In this work, a synthetic matrix was constructed to simulate one year of daily measurements of 39 PM2.5 species concentrations. This matrix, along with 500 replicate datasets created by a balanced bootstrap method, was analyzed by a PMF model and the solutions were classified by neural networks based on chemical profiles. The differences between modeled time series and corresponding input factors used to create the synthetic data give estimates of bias in PMF solutions. The distribution of modeled factors provides estimates on uncertainty in factor contri-

butions.

PMF has been widely used for source apportionment modeling of ambient pollutants. Assessing the uncertainty and bias in PMF modeled factors is an important topic that suits the scope of ACP. This work has been conducted carefully and related discussions are thorough and insightful. The manuscript is overall well organized and very well written. I recommend publication after the authors attend to following comments.

1. The PMF results shown in Fig. 2 and 3 are not very encouraging. This model seems to have trouble retrieving the source profiles and contributions of all the input factors. The differences between modeled and input results are so large for some factors, like fac 6 (gasoline vehicles), fac 2 (vegetative detritus) and fac 8 (meat cooking), that it makes one worry about the usefulness of these source apportionment results, although factors of ammonium sulfate and nitrate and diesel vehicles are well modeled. It is a question how good the PMF modeling was done in this study. Whether some of the input source profiles and/or time series are too closely correlated? How representative was the error matrix? Is fpeak = 0 the best choice? I suggest more details and discussion are given on the PMF modeling work done in this study. It will also be useful if the correlation coefficients between individual source profiles and species time series are given in some way.

2. The behaviors of Fig. 3c and 3h are a bit strange – the bootstrapped time series are always larger than the base case time series (blue $>$ black).

3. p 2984, line 21-24, is it a good assumption of O = 0.3*OC? It is known that the O content of ambient organic aerosols is a function of organic aerosol composition. O:C ratio is usually higher in summer than winter due to enhanced SOA production. Besides, keeping O the same temporal pattern as OC may give OC additional weights in the PMF analysis.

4. In the Discussion, the authors suggested that 4 of their factors are well-modeled and that the not well-modeled factors, including those corresponding to vegetative detritus,

ACPD

8, S1607–S1609, 2008

natural gas combustion, gasoline vehicles, wood combustion and meat cooking, might represent pollution sources for which the PMF model is suspect to provide bias or generally poor fit. I suggest such comments to be carefully qualified. These results are likely only relevant to this study. There are more fundamental reasons for the irretrievability of some factors by PMF, or other multivariate models.

5. Kurtosis is not a commonly reported statistic parameter. It will help to give definition and briefly discuss how it is computed.

6. Extra "puted" on last line, p2981

7. Fig. 2a and 2b seem to be mislabeled. According to the species order given in Table 2, Fig 2a should be for Ammonium Nitrate and Fig. 2b for ammonium sulfate.

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 2977, 2008.