

Interactive
Comment

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

Anonymous Referee #3

Received and published: 17 May 2008

The paper presents a new method to assess the uncertainty of the Positive Matrix Factorization (PMF) source-receptor model application in aerosol studies. The method represents a significant step forward. The comparison of its results with known solutions for a synthetic dataset provides useful insights about the behavior of PMF. The paper does a good job of presenting the approach and its testing but is lacking in the analysis of the results and their implications. The estimated uncertainty is often significantly smaller than the actual one and it is not clear how to take advantage of these improvements.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper



Major issues:

(1) A number of potentially important sources of aerosol is missing: secondary organic aerosol, fugitive dust, non-catalytic engines, a variety of industrial sources, etc. While I do understand that this is just an experiment, the authors should discuss the implications that the addition of the above sources in the synthetic dataset on their conclusions. Some of them have the potential to create significant problems for PMF because there are no tracers for them in the dataset (secondary organic aerosol), they do look like one of the sources (dust), may be small but may be very rich in one of the tracers (non-catalytic engines), or may appear as spikes in the dataset (industrial plumes).

(2) The discussion of the factor contribution plots should be improved. The authors conclude that the presented method could also serve as a way of qualifying future PMF solutions. It is not at all clear how this could be done when the correct answer is not known. For example, how could somebody understand that the contribution of factor 2 is biased but has excellent correlation with the correct solution? Or the behavior of the contribution of any other factor?

(3) The present method of estimating the uncertainty of the contributions of the different sources often fails to capture the true uncertainty. This is a serious problem and should be stressed both in the abstract and the conclusions. Also some of the potential reasons for this failure should be discussed. Given the idealized environment of application (synthetic dataset) and that the behavior should be worse with the real world data, this work appears to suggest that it will be very difficult if not impossible to use such approaches to epidemiological applications. Some discussion is obviously needed.

(4) The paper misses a conclusions section. For example, it is not clear if the method proposed here meets their original objective that is to provide better measures of uncertainty associated with the contribution of specific sources to fine PM.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper



Other issues:

(5) The authors conclude in the abstract that the results are likely dependent on characteristics of data. However, they provide very little information about the synthetic dataset that they have created. They do reference other work, but there should be enough information in the paper to help the reader assess these characteristics.

(6) What does the method used to create the synthetic data set assume about the spatial distribution of sources? One of the major challenges in air quality receptor modeling is the correlation between source contributions located in the same direction (e.g., major roadways for diesel and automobile). They are often difficult to distinguish.

(7) The assumed measurement error and detection limits in the 39 synthetic species should be included in the table.

(8) The benefit of using neural networks (instead of linear correlation) to align factors between PMF solutions should be quantified if possible. It is a major feature of the proposed approach and it not clear how much it helps given the required effort.

(9) I found the order of presentation of the material a little confusing. The novel aspects of the method are presented before the fundamentals of PMF. Bootstrapping is discussed early on, but is explained in the middle of the paper (section 2.3).

(10) A reference to PMF2 is needed.

(11) Near the end of section 2.1, the paper stresses that all solutions are based upon nine factors. Later on, solutions with eight factors are also discussed.

(12) The choice of FPEAK is usually important for the results of PMF. Do the conclusions of the present paper depend strongly on the choice of FPEAK=0? Some discussion is needed.

(13) The point about the need to include oxygen in the list of the synthetic species for mass closure purposes is not entirely clear. Given that the synthetic species represent

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper



a small fraction of organic aerosol, mass closure cannot be achieved in any case.

(14) I am rather skeptical about the inclusion of ammonium sulfate and ammonium nitrate in this type of analysis. The method cannot provide the sources of these secondary species, additional factors are needed and introduced and often the solution (see figure 3) is poorer than just assigning the measured species to the specific sources.

(15) The need to include ammonium is also not clear given that the assumed sources are ammonium nitrate and ammonium sulfate. Its role would be more interesting if there was ammonium bisulfate in the system.

(16) The legends in Figure 2 (base case and actual profile) are misleading. The lines in the diagram are not dashed and there are a lot more than two.

(17) Some discussion of the source profiles created for training the neural network is needed. Some of them have significant differences with each other and the actual profile while others do not. What is causing this behavior?

(18) The results of the seven and ten factor solutions discussed by the authors could be added to those in Table 3 for completeness.

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

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

