Interactive comment on “Chemometric analysis of aerosol mass spectra: exploratory methods to extract and classify anthropogenic aerosol chemotypes” by Mikko Äijälä et al.

Anonymous Referee #2

Received and published: 11 October 2016

This manuscript describes an interesting application of cluster analysis for analysis of ambient aerosol data obtained with an Aerosol Mass Spectrometer. In this method, short pollution time periods are analyzed with positive matrix factorization. The factorization yields background and pollution factor mass spectra that are then analyzed with cluster analysis to classify the distinct types of pollution factors that are obtained. As currently written, this manuscript introduces technical details of an analysis method and would seem to be more appropriate for AMT than ACP. It is important for the authors to highlight how this technique provides improved or new insight into study of atmospheric aerosols so that inclusion in ACP is better justified. I recommend publication in ACP after this change and changes suggested below are made:

We thank the referee for his/her valuable comments.

In accordance with similar comments by referee 1, the manuscript title was modified, large parts of the abstract, introduction, and conclusions were rewritten, and emphasis on the aerosol chemical conclusions was added.

Main Comments

1) In general, the paper is a little longer than it needs to be because it includes a lot of detailed background information about some topics while not enough information is given about necessary details. For example, page 8, Lines 10-13 only offer no quantitative information about how air pollution events are selected. Can some of the words such as “temporary”, “distinct rise”, and “unambiguous separation of pollution plume from background” be quantified?

We have attempted to implement most of the changes proposed to include all necessary details.

We have added the estimated thresholds used in the manual event selection along with more detailed descriptions for our selection criteria in Sect. 2.2.2. However, since the examination was only by visual inspection, and not a mathematical one, the limits are approximates.

We agree with the referee on that exact, quantitative criteria for pollution event selection would certainly be desirable and preferential to the approximated thresholds we used. This would eliminate on of the final sources of subjective judgement of this work. However, in this study, the selection was based on the said criteria due to our difficulty of properly evaluating the fulfillment of especially criteria 2 and 3 (page 8) on an exact level.

2) In figure 1, pollution events of varying time scales and multiple apparent pollution peaks are seen. What exactly is the process used to make these selections?

This is covered in previous comment/answer (#1). The pollution episodes are selected based on the criteria mentioned above. The selection process description is now expanded on in the text.

What controls the length of the time period that is used as a pollution event?

The time of “increased OA concentration”, as verified by a PMF factor emerging and disappearing in the spectra extraction analysis.

Is wind direction data used for selection?

Wind data is not used in selection of pollution events, as there are events that could arise from momentary emissions without a change in wind direction (e.g. passing vehicles, cooking, igniting a fire at a fireplace).

What is the sensitivity of the PMF pollution event solutions to the exact time period range selected around the pollution event?

In our experience the solutions are robust once the time window is kept “short enough”, so that the variability exhibited by the plume/episode forms a major part of the total variability in aerosol mass during the particular period. In this case the solutions are not sensitive to changes in the exact time window selection.

Extending the time window to longer periods, where the variability arising from other reasons (biogenic SV-OOA diurnal cycle, other, consecutive or partly overlapping pollution episodes from different sources, etc) starts to overly dominate the solution, does degrade the solutions quality. Typically this happens when extending the time window to several days (assuming a pollution episode of few hours). Ulbrich et al. (2009) estimate this variability “limit of detection” of separating a factor to be 5% of total variation in OA mass. We find their estimate agrees with our experience with the PMF runs of this work.

Did you consider as an alternative to this manual plume method to run a traditional PMF analysis on the entire dataset and identify plumes as time periods where the residuals of the PMF analysis are high?

The idea of looking at PMF residuals is an interesting one. We did not test it. However, it is a different philosophy in itself, as it pinpoints the mass/variation unexplained by the “standard” PMF model solution chosen. Some issues with this approach are:
As the unexplained mass/variation fundamentally does not equal a separate pollution (or even aerosol) type/source, but might equally well have to do with e.g. volatilization/condensation of semivolatiles, or, oxidation changing the composition of OA over time, or derive from a technical issue such as bad uncertainty estimate.
The definition of “traditional PMF” varies considerably and the analysis is usually at least as subjective as our current methodology, due to similar manual selection of correct solutions, but with fewer pointers to what would be the optimum solution.
Considering the close similarity of e.g. biogenic background SV-OOA and the sawmill SOA pollution, the two would likely get combined in one factor in “traditional PMF”, producing only small residuals, and thus likely missing this important source. Likewise for the oxidized aerosol types (biogenic LV-OOA vs A-LV-OOA).
Residuals are additionally produced by violations of PMF’s underlying assumptions, mainly the idea that factor mass spectral profiles are constant over time (e.g. Ulbrich et al., 2009) – this is often unrealistic for atmospheric aerosols, and the issue is exacerbated for long time series, presumably exhibiting more chemical processes.
To combat these effects were key reasons to split the PMF examination to small time windows in the first place. Using shorter timeframes model assumptions of constant factor profiles can be better assumed to hold, and the separation be more likely driven by actual source based separation than reflecting ongoing chemistry. In short, we foresee trying to pinpoint pollution episodes/types from the residuals would very likely run aground with even worse demarcation issues and subjectivity problems than the current OA mass based selection.

Discussion on these issues was added also to Sect. 2.3.1.

3) It is not clear to me why the pollution event PMF analysis used in this manuscript necessarily provides a more unambiguous separation of “pollution” and “background” than PMF analyses that are performed on the whole dataset. In fact, if the pollution event is simply a result of changes in wind direction that mixes in a different well mixed airmasses, then the PMF factor that is extracted would necessarily be just an average mass spectrum of all the sources present in the polluting airmass. No advantage would have been gained by this method to allow separation of individual sources and this would seem to be a weakness of this method. This aspect is not discussed in the manuscript.

This discussion also links to the earlier question and its answers.

If two sources of pollution are (practically) collocated, unconstrained PMF is unable to differentiate between the sources. This is an inherent limitation of the said data reduction method / receptor model (whether doing the entire data or an episode). In this case the sources are attributed to the same factor (“average mass spectrum”).

However, assuming we are not only relying on one single source of a specific aerosol type, which we believe to be the case given e.g. the variable wind direction distributions (supplementary Figure S.11.), we would expect not all the emission sources will be similarly collocated (and emissions produced in in similar fractions), and thus some sources would produce a more “pure” sample of e.g. HOA or COA. This should be especially true for the close-by sources of e.g. passing cars or emissions from the forestry station. The collocated sources would then be expected show up between the pure samples, which is what we hypothesise to be happening with the weak (transported aerosols, A-SV-OOA) clusters.

We agree that this discussion on mixed pollution events is needed and have added it to the manuscript, to Sections 2.3.1, 3.4 and the conclusions.

Running PMF over the whole data set would not capture the plumes due to the 5% “limit of detection” mentioned earlier. Similarly, it would not provide the separated spectra needed for the clustering applied here. Finally, and our approach is also more robust when it comes to selecting optimal PMF solutions, as described in sections 2.2.2 and 2.3.1. We fix the rotation using “external” information on factor time-series, i.e. what a physically correct (albeit qualitative) description of time series behaviour is like.

4) The manuscript refers to ambiguities in PMF analysis as a weakness and implies that this analysis somehow solves or provides a better solution to this problem of ambiguity. In fact, the manuscript clearly states the difficulty of separating the various primary aerosol sources.

See previous answer.

Regarding the second part, the difficulty of separating primary sources spectra (from each other), discussed specifically in Sect. 3.4.2 is not connected to the question of rotational ambiguity of the PMF model, but the performance of distance metric used in clustering.

One of the advantages of the traditional method of doing PMF or ME- 2 on the entire dataset in this context could be the fact that it can exploit differences in temporal profiles of primary sources (i.e. different diurnal cycles) and also exploit the fact that source mass spectra are similar to allow for separation of multiple primary sources within a well mixed pollution event (i.e. a event such as that mentioned in comment 3 above).

As we can see from the Supplement Fig. S.11 and S.12, the pollution observations are very limited in number, and do not amount anywhere close to forming statistically relevant diurnal patterns

(partly also due to transport times). For a remote station such as SMEAR II, diurnality analysis of pollution plumes (whether by PMF or other receptor models), thus seems unfeasible.

Our intention is in no way to discredit PMF or constrained ME-2 itself, but merely to provide a robust, experimental basis for e.g. using objective and realistic constraints for an improved supervised analysis (e.g. constrained factors).

A comparison between the classification results and a traditional PMF of the entire dataset would have been a good way to address this and to highlight similarities and differences in results. The manuscript should more clearly state discuss the advantages/disadvantages of using this method compared to PMF.

A "traditional" (constrained) ME-2 analysis of two of the data sets ("March 2009" and "September 2008") has been published by Crippa et al. (2014), and is referenced in Sect 2.2.1. In their work Crippa et al. separated LV and SV OOA components, and additionally constrained BBOA and HOA using a reference spectrum (from Paris). In our work we compare our results against these reference spectra (described in Crippa et al., 2013), and find the similarities very high (Sect 3.4). A traditional PMF has been published by Corrigan et al., 2013, but it only managed to separate a BBOA factor besides the general SV and LV components. E.g. Canonaco et al., (2013) and Crippa et al., (2014), highlight the difficulties of separating primary sources for a rural background station like ours in an unconstrained analysis, undermining the feasibility of extracting a wide range of primary OA factors.

As suggested by referee 1, we have modified this section to focus on what scientific insights our approach offers in addition to PMF, rather than implying a "competition" between the two approaches.

5) The strongest part of this manuscript is the application and interpretation of the various clustering metrics to understand similarity and differences between the cluster spectra. It may be useful to highlight more strongly how these metrics could be applied to spectra obtained with typical PMF/ME-2 analysis. Would use of the cluster analysis metrics to reference spectra and PMF solutions provide a means of automating classification in PMF analysis? Also an intriguing part of this that could be discussed in more detail is the possibility to use the cluster analysis to define a-values and reference spectra for ME-2.

The prospect of automatic classification of PMF results is in our opinion definitely a feasible one, and indeed one of the motivations for this type of a study. The application could be, for example, in classifying a large set of bootstrapped PMF runs or a large number of PMF "seed" runs or Fpeak runs' results. Similarly, we hope the spectral similarity metric optimization would be useful for a) evaluation of PMF result factors' similarities against each other and b) identification of PMF result spectra (against library references, such as the AMS Spectral Database). We discuss the application of cluster centroid spectra and within-cluster variation as fitting input for a constrained ME-2 analysis (in the Introduction, Sect 3.6 and Conclusions), but have now highlighted this further, and added a paragraph in conclusions summarizing these future prospects. We thank the referee for these good suggestions, and have added emphasis to them.

---

references:

Canonaco, F., Crippa, M., Slowik, J., Baltensperger, U., and Prévôt, A.: SoFi, an IGOR-based interface for the efficient use of the generalized multilinear engine (ME-2) for the source apportionment: ME-2 application to aerosol mass spectrometer data, *Atmospheric Measurement Techniques*, 6, 3649-3661, 2013.

Crippa, M., DeCarlo, P., Slowik, J., Mohr, C., Heringa, M., Chirico, R., Poulain, L., Freutel, F., Sciare, J., and Cozic, J.: Wintertime aerosol chemical composition and source apportionment of the organic fraction in the metropolitan area of Paris, *Atmospheric Chemistry and Physics*, 13, 961-981, 2013.

Crippa, M., Canonaco, F., Lanz, V., Äijälä, M., Allan, J., Carbone, S., Capes, G., Ceburnis, D., Dall'Osto, M., and Day, D.: Organic aerosol components derived from 25 AMS data sets across Europe using a consistent ME-2 based source apportionment approach, *Atmospheric chemistry and physics*, 14, 6159-6176, 2014.

Corrigan, A., Russell, L., Takahama, S., Äijälä, M., Ehn, M., Junninen, H., Rinne, J., Petäjä, T., Kulmala, M., and Vogel, A.: Biogenic and biomass burning organic aerosol in a boreal forest at Hyytiälä, Finland, during HUMPPA-COPEC 2010, *Atmospheric Chemistry and Physics*, 13, 12233-12256, 2013.

Henry R.C.: Current factor analysis models are ill-posed. *Atmospheric Environment*, 21, 1815-1820, 1987.

Paatero, P.: Least squares formulation of robust non-negative factor analysis, *Chemometrics and intelligent laboratory systems*, 37, 23-35, 1997.

Paatero, P., Hopke, P. K., Song, X.-H., and Ramadan, Z.: Understanding and controlling rotations in factor analytic models, *Chemometrics and intelligent laboratory systems*, 60, 253-264, 2002.

Paatero, P., Eberly, S., Brown, S. G., and Norris, G. A.: Methods for estimating uncertainty in factor analytic solutions, *Atmos. Meas. Tech.*, 7, 781-797, 2014.

Ulbrich, I., Canagaratna, M., Zhang, Q., Worsnop, D., and Jimenez, J.: Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, *Atmospheric Chemistry and Physics*, 9, 2891-2918, 2009.