

Interactive comment on “Positive Matrix Factorization of Organic Aerosol: Insights from a Chemical Transport Model” by Anthoula D. Drosatou et al.

Anthoula D. Drosatou et al.

spyros@chemeng.upatras.gr

Received and published: 19 November 2018

(1) Drosatou et al. present a study in which chemical transport model predictions of organic aerosol over Europe are used in a PMF (and ME-2) analysis to determine what types of organic aerosol factors exist in model predictions. The use of a chemical transport model means that the PMF factors can be directly examined in terms of their sources and identity. The major findings include the composition of POA factors in terms of potential SOA contributions and expected error as well as identification of 2 types of SOA or OOA-like factors in a variety of locations. They show that the two types of SOA do not separate anthropogenic and biogenic SOA and the separation is mainly

C1

based on age. This is a useful analysis and comments below focus on two main areas.

We appreciate the comments and suggestions of the reviewer. Our responses and corresponding changes in the manuscript (in regular font) can be found below after each comment (in italics).

Main comments:

(2) How would results be different if source information (for example the chemical identity of POA and bbPOA) was not used in the PMF analysis of model output? If the volatility of the model-predicted OA was the only chemical information in the PMF analysis, would you get similar results? This may provide insight into how results from this work translate to AMS analysis in which sources may not be very chemically distinct due to fragmentation.

This is an interesting suggestion allowing the PMF to focus just on the volatility of the OA. We have repeated the PMF analysis using only the volatility distributions. We first tried two factors. The corresponding PMF factors included material from all volatility bins. So PMF did not separate the OA into semi-volatile and low volatility material. In the next step we assumed three factors but still the factors included surrogate compounds with a mixture of volatilities. We have added a new sub-section in the Results to discuss the results of this test simulation that further supports our conclusion that PMF does not separate OA components solely based on their volatility.

(3) PMF appears to have been performed on a site by site basis. Can this be clarified? The analysis generally always leads to two SOA (or OOA) factors, but the composition of the SOA factors varies by site. How many SOA factors would be obtained if all locations from the model were used in one PMF analysis? An analysis along these lines could help inform questions in the second main comment regarding how different the two OOA factors are in different locations or studies.

C2

Indeed, the analysis presented in the original paper was performed for each site separately similarly to the analysis of field campaign measurements. This is now clarified in the manuscript. We have followed the reviewer's suggestion and performed an additional test in which we combined results from all seven sites (all the examined areas in the paper). The application of PMF to this comprehensive set resulted in four factors: fresh biomass burning, other primary OA and two secondary OA factors (fresh and aged SOA). The number and character of the factors were similar with the site by site analysis, but there were differences in the composition and contribution of the factors. The results of this test are now discussed in the revised paper.

(4) *What information is introduced by the PMF analysis of model output that is not otherwise available? Could the same "factors" be obtained by determining how much POA, SOA-sv, bSOA, aSOA, etc correlate or covary and making two groups?*

The reviewer is correct; factor analysis methods are in general based on the temporal correlation among the concentrations of different pollutants. However, in their effort to limit the dimensionality of the chemical (or AMS m/z) space, these approaches distribute the pollutants into factors in ways that are by no means transparent. Our goal in this work has been to shed a little more light on what PMF does when it is applied to the AMS organic aerosol data. We have added this brief discussion in the paper.

(5) *How was the boundary condition OA at the edge of the domain specified and evaluated? If the boundary OA was not assigned a C^* of $0.01 \mu\text{g}/\text{m}^3$, would it have ended up in a different factor?*

In general boundary conditions of regional chemical transport models are obtained from the output of similar global models or from some averages of measurements. For the purposes of this work we assumed low volatility constant OA boundary conditions. This choice facilitates the separation of highly aged OA coming from outside the modeling domain from the fresher material that is emitted and/or produced inside the

C3

domain. Obviously, the absolute OA concentrations especially near the boundaries of the domain can be dominated by these boundary conditions. To avoid such issues, we have used sites that are far from the boundaries. Overall, our conclusions are quite robust to the choice of the OA boundary condition values. The effect of our choices of OA boundary conditions is now discussed in the paper.

(6) *Meaning of two SOA or OOA factors. Is the proliferation of terms in literature (LO/MO-OOA, OOA-1/2, LV/SV-OOA) indicating true site to site variability in the OOA components or is it just a nomenclature choice?*

We believe that the evolution of the terms used to describe these two factors reflects our understanding (or lack there-of) of the nature of these factors and not so much site to site variability. The use of OOA1 and OOA2 reflected the complete lack of understanding. Then the use of Less and More Volatile OOA showed the beginning of some understanding, but it has probably led to some confusion and a few misconceptions. The next step (use of Less and More Oxidized OA) is probably more accurate. Our work here supports the hypothesis that these factors correspond to Less and More Aged OOA present in each site. We have added this discussion to the corresponding section of the paper.

(7) *The authors argue that the designation of the two AMS factors based on volatility is somewhat misleading due to overlap in their volatility. I was not convinced that this designation was misleading (based on Figure 10) but do agree that it is a simplification. What is the best description of the two factors given that they likely overlap on many metrics (volatility, O:C, age, etc) and age, O:C, and volatility covary?*

The use of the volatility-based terminology suggested to most of us that there is a volatility threshold and OA components that are more volatile than this are grouped by PMF in one factor (e.g., SV-OOA) and the less volatile compounds in the second (LV-OOA). Our results both from this theoretical analysis but also from direct volatility

C4

measurements of AMS factors (Paciga et al., 2016; Louvaris et al., 2017) suggest that this is not the case. The so-called semivolatile factor may include very low volatility OA and vice versa the so-called low-volatility factor may include semivolatile material. We believe that the use of more and less oxygenated is safer and that the use of more and less aged will be probably proven to be more accurate.

(8) *How should CTMs evaluate their predictions compared to AMS data beyond SOA vs OOA? Can the analysis here be used to provide a range of agreement where models can be assumed “in agreement?”*

This is an excellent question. Our results suggest that the comparison of CTM predictions of POA and fresh biomass burning OA to the corresponding AMS results is meaningful if these are major sources and taking into account the uncertainties estimated here. The comparison of the less and more volatile OA predicted by CTMs to the corresponding OOA factors is probably not a good idea. Summation of the two OOA factors into just OOA appears to be quite safe based on our results here. On the other hand, if a CTM can keep track of the age of OA the comparison of more and less aged predicted OA to the two OOA factors could be potentially useful. We have added this discussion about model evaluation to the paper.

Minor comments:

(9) *Lines 22-28 of the abstract are useful, but could be condensed. Mentioning the fraction of the POA factor that is secondary (e.g. lines 477-478) would be even more useful.*

We have followed the reviewer's suggestion and rewrote this part of the abstract.

(10) *Lines 81-104 are missing the MO- and LO-OOA designations (Xu et al. 2015 PNAS <https://doi.org/10.1073/pnas.1417609112>) in the discussion and how those fit with the other AMS PMF factors from literature.*

C5

We have tried to keep the terminology used in the corresponding studies to both show the evolution of the nomenclature but also the rather confusing picture. We have added a sentence indicating the correspondence of the terms (for example OOA-1, LV-OOA and MO-OOA have been used for the same factor in most studies).

(11) *Line 134: what version of SAPRC was used?*

A version of SAPRC99 extended to include the VBS species was used in this work. This information has been added to the text.

(12) *Paragraph starting at line 250: Clarify that there was no observed POA or bbPOA factor in observations or model for Melpitz.*

We have added this clarification in the revised paper.

(13) *Figure 3: Could boundary OA and POA+S/IVOC-SOA be added to panel (b)? Could SOA (excluding boundary and S/IVOC-SOA) be added to panel (a)? How much value does the PMF factor bring compared to classifying boundary and S/IVOC-SOA as one type and all other SOA as one type?*

The major point of this figure and the corresponding example of the PMF analysis in Melpitz is that the analysis of the ambient AMS dataset and that of the PMCAMx VBS predictions results in the same number of factors with a very similar diurnally-averaged behavior. This supports our hypothesis that the PMF analysis of the PMCAMx predictions can help us understand better the results of the PMF/AMS analysis. One can try different combinations of the predictions of PMCAMx and compare them to the results of PMF/AMS (see for example Fountoukis et al., 2014 for such an effort), but this is outside the scope of this work. Please note that we are not suggesting that PMF should be applied to CTM results. CTMs (as we show in this work too) provide directly information about sources of OA without the need of PMF. A brief discussion of this point has been added.

C6

(14) *Figure 7: How were the locations chosen?*

We have added a description of our criteria for the choice of these locations. Briefly, Majkow Duzy (Poland) has the highest predicted contribution of POA to OA. St. Petersburg, Catania and Majden are three locations in different environments with high bbOA levels during the simulation period. Melpitz, Cabauw and Finokalia were chosen because there are AMS measurements available for the simulation period and they also cover quite different environments. The other sites in Figure 7 were chosen because they had different predicted bbOA/levels and they could help us get information for the full range of values.

(15) *Figure 7: What would you expect the NME to be for typical urban, suburban, or rural conditions (add to plot)?*

Based on our analysis the actual contribution of bbOA to the total OA is more important than the type of environment for Europe at least. If a site (even if it is urban) is influenced by a major nearby fire contributing significantly to the OA then PMF does well in quantifying its impact. If on the other hand if a rural site is only marginally affected by far away fires then the corresponding normalized error can be significant.

(16) *Did the model include any aqueous SOA? Where would that appear in the PMF analysis?*

No, this version of PMCAM did not include aqueous SOA production. The treatment by PMF of such OA that has been produced by a different pathway is an interesting question for future work.

Editorial comments

(17) *Line 126: "a regional three-dimensional CTM" is duplicated within the sentence.*

We have rewritten this sentence.

C7

(18) *Figure 1: Needs (a) and (b) labels or titles.*

We have added the corresponding labels.

(19) *Figure 2: Could be on same panel in different colors.*

We would prefer to keep them separate. The figure becomes rather confusing when these two timeseries are in the same panel.

(20) *Figure 4: Could go in SI.*

We have followed the reviewer's suggestion and moved this figure to the SI.

(21) *Figure 11: switch columns 3 and 4 so that both Melpitz and Finokalia data reads as increasing age going left to right.*

We have switched the order of the two Finokalia factors for consistency with the Melpitz factors.

Interactive comment on Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2018-802>, 2018.

C8