

## ***Interactive comment on “Volatility of organic aerosol and its components in the Megacity of Paris” by A. Paciga et al.***

**A. Paciga et al.**

spyros@chemeng.upatras.gr

Received and published: 15 January 2016

*(1) The authors report results from measurements of organic aerosol volatility, as characterized through measurement of the loss of OA mass due to evaporation as a function of temperature, for ambient particles sampled during winter and summer in Paris, France. They present results both for the total OA behavior, as well as for the behavior of specific OA “factors” determined from positive matrix factorization. They ultimately find that the average OA volatility is not all that different between winter and summer in Paris, but that there are distinct differences in the OA factor volatilities within a given season and between the same factors in different seasons. These differences are discussed briefly in terms of differences in chemical composition, as reflected from likely source attributes and obtained mass spectra of the OA factors. Overall, this study con-*

C11608

*tributes new information regarding the volatility of ambient OA, but it could do a better job of putting their specific results in the context of literature results. Additionally, I believe that the data associated with the OA factor volatilities could be presented in a more tangible way to the reader, and suggestions are provided as to how to do this. I have some concerns regarding the averaging procedures used in terms of how they may/may not introduce any particular biases into the analysis. I suggest that this work is ultimately publishable, but should be considered further after revision.*

We appreciate the constructive assessment of our paper by the referee. These individual comments are addressed below.

*Specific comments:*

*(2) The authors report measurements of the mass fraction remaining, which was determined from measurements made alternately every 5 minutes sampling through the TD or ambient particles. However, it is not made clear exactly how the ratios were calculated. Were they calculated using just measurement pairs (e.g. the measurement at t-1 divided by the measurement at t) or were they calculated using an average of TD measurements made before and after each ambient measurement? Or an average of ambient measurements made before and after each TD measurement? To some extent, all of these details may cancel out after sufficient averaging, but the normalization process can certainly impact the apparent amount of atmospheric variability, as shown in Fig. S1. The normalization procedure is less of a concern when atmospheric conditions are constant, but when things are changing rapidly (and here, rapidly means on 5-15 minute time scales) the method used can matter a lot. Can some of the “atmospheric variability” in Fig. S1 be explained away as a result of the normalization procedure used? Have the authors considered, for example, filtering their measurements based on the extent of change between two sequential ambient points? I believe that such considerations are particularly important in the splitting of the data set into high/low concentration periods since visual examination of Fig. 1 suggests that there is*

C11609

*generally greater atmospheric variability (faster changes) during the high concentration periods than during the low concentration periods. Ultimately, I believe that the authors need to at minimum be more specific as to how their MFR values were calculated.*

The MFR values were first calculated using measurement pairs (the measurement at interval  $i$  divided by the measurement at interval  $i+1$ ) and then they were further averaged resulting in Figure 3. The final averages were relatively insensitive, as the reviewer guessed, to the details of the individual MFR calculation. However, the variability of the atmospheric conditions, once more as the reviewer suggests does introduce variability in the individual MFR estimates. We had the same concern as the referee regarding the variability of the MFR during the high concentration periods and this was one of the motivations for splitting the data set into high and low concentration periods. However our analysis did not find any differences for the corresponding MFRs during these two types of periods. We have added in the revised manuscript the information about the details of the calculations of the MFR and the variability that is introduced due to the variation of the atmospheric concentrations.

**(3)** *P22267, L25: The authors mention the flows for the SMPS, stating at the end that the values selected “: : extend[ed] the size-range of measured particle: : :”. Extended relative to what? Had the SMPS been operated at 3 lpm sheath, 0.3 lpm sample, the range would have been “extended” even further? In other words, I don’t find this statement necessary nor clear.*

We have revised the corresponding sentence simply stating the operational parameters of the SMPS during these measurements to avoid unnecessary confusion.

**(4)** *P22268, L26 and Fig. S1: The authors note the “experimental variability” and use this as a reason to not split the data set. I believe that this is reasonable, but at the same time it would seem to me that Fig. S1 could be presented as, for example, a box and whisker graph or as means and medians to help illustrate whether the average (or*

C11610

*median) properties differed between the high/low periods. There is currently so much atmospheric variability that it is difficult (for me at least) to really conclude that there is no “discernable difference.”*

We do agree with the suggestion of the referee. In the revised paper we present the data in Figures S1 and S2 as box and whisker plots. The new Figure S1 does illustrate the similarity of the two data sub-sets a lot better.

**(5)** *Section 2.2: The authors have ultimately averaged their data into 5 degC bins. It is not clear to me what guided this decision, as it is clear in looking at the rawer data (Fig. S1 and S2) that the data points are not equally distributed with respect to temperature. There are clearly a few temperatures that were favored, i.e. have lots of points, and many temperatures at which there are only a few data points. For example, it appears that the 100 degC bin might have 100’s of points (or at least many 10s of points), whereas the 80 degC bin will only have a handful of points that make up the average. This means that the sampling is not equally weighted at all temperatures with respect to sampling period. For the winter data, such potential statistical issues visually look like they might not be particularly important, since they data show less overall variability compared to the summer data, where the variability can be quite large. Further, it appears that the TD operation was different between summer and winter, and thus the number of points sampled at especially the higher temperatures is quite different between the two studies, with (for example) many points at 150 degC for winter but very few in that same range for summer. I believe that the authors need to justify their 5degC bin choice and would actually recommend that they actually consider wider bins. Even 10 degC would be more justifiable than 5 degC, in my opinion. In any case, given that they have apparently binned their data into 5 degC bins, I am somewhat confused by the data presented in Fig. S3, for which the temperature points seem to be almost randomly distributed. But, perhaps this is simply an issue of there being some “missing” temperature ranges (where there was insufficient data?) that makes it seem like there is not a point every 5 degC. However, this still wouldn’t explain why,*

C11611

specifically, the LVOOA thermogram has so many fewer points than any of the other thermograms given that the “percentage of measurements above threshold” for this factor is equal to or greater than those for HOA and COA, and very similar to that for MOA.

Our choice of the 5 C as the averaging interval was a compromise between averaging enough points at similar conditions and maintaining the dynamic behavior of the thermogram. While the choice of 10 C suggested by the reviewer is also reasonable, the corresponding analysis does not add anything new to the results or our conclusions. There were indeed different numbers of data points in each bin. This is taken into account into the calculation of the standard deviation of the mean for each temperature bin. During the first campaign we had a lot more measurements in specific intervals (e.g. from 80-100 C) in an effort to learn more from the corresponding variability. The corresponding efforts did not result in any significant insights and therefore during the second campaign we distributed the measurements more uniformly in the various temperature ranges. The missing points of the LV-OOA are from temperature bins with few measurements and very large corresponding uncertainties. We have added explanations of these points in the revised manuscript.

**(6)** *It is indicated that PMF analysis was performed “combining both ambient and thermodenuded spectra” with a reference to Crippa et al. (2013) given. However, in reading Crippa et al., although it is noted that the AMS was operated with the TD, it is not clear whether the PMF analysis performed in that study was performed using the TD data as well. I suspect it was not. That would mean that the PMF analysis mentioned here is a new product, and as such the key aspects of the PMF analysis should be provided (perhaps as supplementary material), such as how the number of factors were selected, uncertainty estimation, etc.. Was the analysis performed here independent, or guided by the results of Crippa et al. already in terms of e.g. the number of factors obtained?*

C11612

The analysis presented in Crippa et al. (2013) was performed using only the ambient measurements. This analysis was repeated for the purposes of this work combining the ambient and TD datasets. This second analysis was guided by the original results (e.g. the same number of factors was used) and produced for all practical purposes the same results for the ambient data for all factors. These are analyzed in great detail in Crippa et al. (2013) and adding the same information in the supplementary information here would not provide any additional insights. This information about the PMF analysis has been added in the paper.

**(7)** *More information should be provided regarding the removal of MFR measurements. The authors state that “to minimize these problems, a minimum ambient mass concentration was determined for each PMF factor, based on the concentration range for which several MFR measurements exceeded significantly unity.” How many is “several”? Are these continuous in time? Or spread throughout the entire campaign? Is this only considered when the TD temperature is greater than some value (such that the MFR would be expected to be much less than one, which may not be the case when temperature is close to ambient, e.g. <60 C). And were any MFR values >1 observed above this criterion? If so, were they included or excluded from this analysis? It is not clear, but given the data shown in Fig. S1 and the extent of “atmospheric variability” for the “Peak Data” I suspect that all MFR >1 have been filtered from the data set. If so, this should be stated clearly. And I assume that the “several” criterion was consistently applied, but this should be stated and, if not consistently applied, justified. Further, it is not clear that the authors have considered the role of atmospheric variability in driving MFR values to be >1 (which is dependent upon the normalization method applied). MFR values greater than unity can be an indication of working below the detection threshold, but they can also reflect the influence of atmospheric variability. The authors could possibly get at this issue by calculating an “effective” MFR value from just the ambient (or TD) measurements, where the “effective” MFR is the ratio between sequential ambient measurements, or the ratio between the average of two ambient measurements that*

C11613

*are around a third and that central value. When this ratio deviates substantially from unity, either "atmospheric variability" or operating near the detection limit are having an outsized influence on the measurements. I say all this in part because it is unclear to me exactly why the threshold concentrations should be different in the winter versus the summer campaign. Did something change about the instrument?*

We agree with the reviewer that additional discussion of the details of the removal of MFR measurements is needed. During low factor concentration periods the estimated MFR values could exceed 2 or they could be practically zero even at low TD temperatures. Also there was significant MFR variability from measurement to measurement. All of these are signs of problems due to the division of two very uncertain values. These problematic MFR values would be during certain periods lasting for a few or sometimes several hours and during all of them the corresponding factor concentration was very low. The factor concentration cut-offs corresponded to the appearance of MFR values exceeding 1.5 (usually a lot more than this). Small exceedances of 1 (say 1.1) were of course observed during measurements at low temperatures and were due to the uncertainty introduced by the estimation of the MFR from the ratio of two consecutive measurements. All the MFR values corresponding to concentrations lower than the thresholds of Table 1 were excluded from the analysis while all the rest (even if they exceeded one) were used. Please note that the winter thresholds correspond to the left tails of the corresponding distributions (they correspond only to a few percent of the values) while the summer thresholds are a lot closer to the middle of the corresponding concentration distributions. These different statistics can probably explain the corresponding small differences in values. Use of the summer thresholds for winter would produce practically the same results in our analysis. These important points are clarified in the revised paper.

**(8)** *P22270, L9: It is not entirely clear what is meant by "For a fair comparison of volatility distributions for these data sets: : :". Fair in what particular way? It seems to me that these values were, to some extent arbitrarily selected. This is fine, as they are*

C11614

*both reasonable. But the "fair" aspect of this selection should be clarified.*

We agree with the point of the reviewer and we have deleted the first part of the sentence.

**(9)** *The authors use the terms SVOCs, LVOCs, and ELVOCs in the abstract, yet these are not defined until P22270 and thus any reader unclear with the specific definitions may find their meaning confusing. I suggest the authors be more precise within the abstract.*

We have added the definitions of these groups of compounds in the revised abstract.

**(10)** *Figure 3: Returning to the 5 degC bins, I find some of the points presented in this figure to be potentially problematic. The authors report three values (with standard deviations!) for the summer campaign in the range 180-200 degC. However, in looking at Fig. S1, it seems clear that there are only 4 total data points in this region. Thus, I do not think it is at all appropriate to show these points in Fig. 3, and certainly not with error bars as it is entirely unclear where these would have come from. Perhaps there is more data than is shown in Fig. S1, in which case perhaps the averaging in this range is appropriate, but then the data presentation in Fig. S1 would be misleading. I strongly suggest that the authors change to using at least 10 degC bins (if not even 15 degC).*

The error bars in two of the data points in Figure 3 were indeed wrong. We do appreciate the correction. We have updated Figure 3 and these points are shown without error bars. As mentioned above the use of the 10 C bins is reasonable but it does not change the results and the conclusions of the paper. For the points in question it also has almost no effect.

**(11)** *Figure 4 and Figure S3: I believe that Fig. S3 must be moved to the main text. The data need to be shown, not just the results from the data fitting. The readers need to*

C11615

more easily be able to assess the differences between the different thermograms and relate them visually to the bar charts shown in Fig. 4. Additionally, it would seem to me that it is possible for the authors to add error bars (e.g. standard deviations) to each of the MFR points for the factors, as these seem to be the binned values. The data are currently shown in their binned, averaged values, which does not give an indication of the variability associated with the factors themselves. I would strongly encourage the authors to show, for each factor, all of the data points along with their binned values and standard deviations. It would additionally be good if the authors considered both the mean and the median values. Are these similar, indicating a normal distribution? Or do they give very different results? This type of presentation of the data would allow the reader to judge the data quality for themselves. With the data as presented, I can only assume that the variability in each of the thermograms for the different factors is similar to that in the total OA MFR. However, I suspect that the actual variability is a lot greater than for the total. But perhaps it is smaller because the authors have filtered out the low concentration points that may contribute disproportionately to the variability in the total OA MFR values shown in Fig. S1 and S2.

We have followed the suggestion of the referee and added the error bars to each of the MFR points for the factors. The corresponding figure does not add much useful information for the average reader of such a paper. It is mostly for the few specialists who do similar type of data analysis. We believe that it fits better in the Supplementary Material and that Figure 3 provides the important information required by the main paper.

**(12)** *Regarding Fig. 4 itself, I strongly suggest that the authors present the volatility distributions using the same x-axis range for all of the figures. As presented, I find that the reader can easily miss the different scales. (And I do not think that just stating in the caption that the scales are different would be sufficient. The axes ranges should be identical.)*

C11616

We have followed the reviewer's suggestion and now use the same x-axes for all of the figures.

**(13)** *Regarding the ranges shown in Fig. 4, it becomes clear that the different OA factors were not fit using the same  $C^*$  range, but instead using an individual range for each one. Otherwise the LV-OOA range is not compatible with the HOA range (for example) if only 6 bins are used. This should be made clear in the text.*

We have clarified this point in the revised paper. Different 6-volatility bin solutions were needed for the fitting of the thermograms of the various factors.

**(14)** *The authors state that their volatility distributions have been obtained "using the uncertainty analysis approach of Karnezi et al. (2014)". It is, however, not clear from the presentation whether this includes experimental uncertainty or, more specifically, if it does include experimental uncertainty how those uncertainty values were established. Standard deviations? See comment above regarding Fig. S3.*

We have rephrased this sentence deleting the rather confusing "uncertainty analysis" descriptor. The approach of Karnezi et al. (2014) finds multiple volatility distributions that fit reasonably well the average values of the thermograms and then weighs them to estimate the predicted volatility distribution and its uncertainty. It does not use the experimental uncertainty.

**(15)** *The concentration filtering method of the authors may potentially lead to sampling biases in terms of the time of day of the main measurements. For example, looking at the average diurnal profile for COA in Crippa et al. (2013) for summer COA, the lowest concentrations tend to be found during the day and the highest at night. Thus, the filtering method will oversample nighttime behavior for summer. Although the same general diurnal profile is obtained for COA in winter, the overall concentrations are much larger and thus fewer points are rejected meaning that the time-of-day sampling bias*

C11617

*imposed by the filtering criterion will not be as strong. To what extent do the authors believe such differences contribute to the apparent seasonal differences in volatility? Related to this, the average spectra considered for each of the various factors have likely not been weighted to be reflective of the periods excluded from MFR analysis. Thus, the spectra may not be fully representative of the actual chemical composition of the particles for which MFR values were determined. Has this been considered?*

This is an interesting point. Our approach implicitly assumes that the AMS spectrum of the factor remains constant during the period of the measurements and that its volatility distribution as a result also does not change. However, there can still be subtle differences that may be hidden. It is very difficult to estimate if there is any bias introduced by our inability to determine the MFR during very low concentration periods. However, we do not believe that this potential bias can explain the significant differences between summer and winter for COA. Our major argument is the observed difference in volatility is consistent with the observed difference in the AMS mass spectra of the factor during the two seasons. Please note that the AMS analysis covers the whole period and no data are excluded.

**(16)** *A comparison with other literature observations is notably absent for the most part. The authors ultimately only compare their derived volatility distribution for one OA factor (BBOA) to the literature. I strongly suggest that the authors compare their work with some of the literature observations that they mentioned in the introduction (p22266). This would facilitate broader understanding of their results.*

We have added comparisons with the volatility distributions of Cappa and Jimenez (2010) for Mexico City as well as the few other such studies in the literature.

**(17)** *There seem to be somewhat “standard” colors associated with presentation of AMS factor data in the literature (including in Crippa et al. (2013)). I encourage the authors to utilize this “standard” color scheme to facilitate easier comparison with the*

C11618

*literature.*

We have done our best with the selection of the colors. Please note that we are also using in Figure 7 different shades to show intensity so we need to be careful to avoid confusion with the use of colors for the various factors.

**(18)** *Figure 6: I suggest that difference spectra might provide easier viewing of the differences in spectra.*

We have constructed the corresponding graphs but on their own they are not that helpful. Given that this is not a paper about PMF analysis of AMS results adding two more figures is not worthwhile. We believe that Figure 6 is sufficient for the purposes of this work.

**(19)** *The last line of the abstract should be removed, as it is vague yet, as it turns out, extremely specific.*

We have deleted the suggested sentence.

**(20)** *I have some difficulty understanding the specific meaning of statements such as “The summer COA was more than one order of magnitude more volatile than the winter COA” (P22276, L11), given that the authors actually present volatility distributions, not single values. What specifically does it mean to state that a distribution is one order of magnitude more volatile than another distribution? That the highest values are different? That the (weighted) mean behavior is different? The authors should be precise.*

We now explain that we are comparing the weighed mean volatility that is also shown in Figure 7 for each factor.

**(21)** *I find the language that the authors “incorporated the results into the 2D-VBS” a*

C11619

*bit imprecise. What does it mean to “incorporate” something into the 2D-VBS? My take on what the authors did was to place their derived volatility distributions and measured bulk average O:C ratios onto a particular graph. But I don’t see how this information has been “incorporated.” Further, the authors should more explicitly recognize the limitation of their placement of their data onto the 2D-VBS framework, specifically that whereas they have volatility distributions they only have bulk average O:C ratios. Thus, they have no information as to how the O:C might have varied within each of the volatility bins for each of the factors. Previous results (e.g. for laboratory SOA, Kostenidou, EST, 2009) suggest that the O:C of a factor may vary with temperature, but not necessarily by a substantial amount. Regardless, this limitation should be noted more specifically, especially in the context of statements from the authors such as “These results indicate that there was not a direct link between the average volatility and the average O:C for these OA components.” They can say this is the case between factors, but they do not present information that this is true within a given factor.*

These are both good points. We have rephrased the sentence and also stressed the limitation of using only the average O:C in this analysis.

---

Interactive comment on Atmos. Chem. Phys. Discuss., 15, 22263, 2015.

C11620