Atmospheric
Chemistry
and Physics
Discussions

Open Access

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

Anonymous Referee #1

Received and published: 21 August 2015

Review of Paciga et al.

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 volatilites 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-

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.

Specific comments:

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 dataset into high/low concentration periods since visual examination of Fig. 1 suggests that there is 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.

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.

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 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."

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 that 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 appoints seem to be almost randomly distributed. But, perhaps this is simply an issue of there being some "missing" temperature ranges (where there was insuffient data?) that makes it seem like there is not a point every 5 degC. However, this still wouldn't explain why, 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.

It is indicated that PMF analysis was performed "combining both ambient and thermod-enuded 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?

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 20 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. <60C). 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 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?

P22270, L9: It is not entirely clear what is meant by "For a fair comparison of volatility distributions for these datasets...". Fair in what particular way? It seems to me that these values were, to some extent arbitrarily selected. This is fine, as they are both reasonable. But the "fair" aspect of this selection should be clarified.

The authors use the terms VOCs, 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.

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

C6132

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).

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 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.

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

C6133

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.)

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 compatable with the HOA range (for example) if only 6 bins are used. This should be made clear in the text.

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.

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 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?

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.

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 literature.

Figure 6: I suggest that difference spectra might provide more easier viewing of the differences in spectra.

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

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.

I find the language that the authors "incorporated the results into the 2D-VBS" a 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, ES&T, 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.

C6136