

Review of Li et al “Predictions of glass transition temperature and viscosity of organic aerosols by volatility distributions”

General Comments

In this manuscript, the authors develop parameterizations of the glass transitions temperature (T_g) based on vapor pressures of a large number of pure organic compounds containing C, H, N, O, and S. The authors compare T_g predictions based on several different algorithms and then use the algorithm to predict the phase of SOA based on SOA volatility measurements from field studies. Finally, the impact of condensed phase water on the aerosol viscosity is examined.

Calculations of the phase of organic aerosol is timely, relevant, and will be of interest to readers of ACP. The number of figures is appropriate and are generally well presented. The writing is clear and generally well organized, though could use some further technical editing.

The authors are attempting to reduce a very challenging problem in the field to a parameterization that can be reasonably included in a global modeling framework with limited data to work with. I applaud them for this effort and fully recognize this is challenging, particularly with the limits in the data. With that said, I think the authors could have done a better job in more clearly addressing the uncertainties and limitations in the inputs to their parameterization early in the manuscript. In addition, large number of parameterizations and different data analysis and parameter estimation methods gets confusing at times. These points can be remedied with relatively minor revisions. Perhaps more challenging is that the results of the parameterization themselves don't seem to constrain the phase of OA particularly well. One example from the text is that for the SOAS campaign the range of T_g for OA is between 232 – 330 K depending on how the same set of data is analyzed. So, a reader (or reviewer) may be either confused or left wondering what exactly these calculations tell us. Adding condensed phase water and its impact on phase/volatility further increased the complexity. I suggest the authors try to better address this uncertainty/error/prediction range more thoroughly. I don't expect they will be able to resolve the issue and I don't have any specific suggestions for how to resolve this, but I think the attempt would significantly improve the impact of the paper. After these corrections, the manuscript would be appropriate for publication in ACP.

Major Specific Comments

Improve clarity of the presentation and take-home message. I had to read the manuscript several times to fully understand the method the authors were using. One challenge is that at least 5 different T_g parameterizations are presented and compared (Eqs. 1, 2, 3-5, 6, and a global model parameterization), and it gets difficult for even a careful reader to keep track them. A few suggestions to improve this: 1) label the figures themselves (i.e. in a title) so the reader doesn't have to study the caption to understand the figure, 2) shorten the captions to be more concise; the key figures have captions that are a paragraph long, 3) refer to the parameterizations by a

descriptive name rather than the equation number (for example, Volatility + O:C, Volatility only, composition only, global model); this is done some places in the manuscript, but not all.

In addition, there are (at least) 4 different methods to infer aerosol volatility, at least 2 methods of “measuring” T_g , and possibly different ways of “measuring” T_m . Again, it got difficult to track what measured as opposed to calculated and what was being compared, given the large number of combinations and permutations of T_g , T_m , volatility, etc. involved. One example: in Figure 3 the x axis is labeled “Measured T_g ” while in Figs 1 and 2 it is labeled “ T_g measured or estimated from T_m ”. Is the data presented in Figure 3 a further subset of the data in Figs 1 and 2? I suspect they are the same data, but this should be made clear.

After reading the manuscript I’m not left with a clear conclusion regarding the utility of these parameterizations to predict the phase of organics in the atmosphere. As the authors point out, T_g predictions span a range of 100 K depending on how volatility is estimated from the same dataset. The predicted T_g range also unfortunately spans the tropospheric T range, so that the uncertainty is precisely in the temperature range where there is high sensitivity to T_g . It wouldn’t matter much if the predicted T_g range was 100-200 K, but that’s not the case. The effect of RH makes this even more unclear, with a very wide range of viscosity predicted below 60-70% RH (Fig 7). For example, line 31 states that T_g varies from 290 – 339 K from which I would conclude that OA should exist predominantly in a glassy state under ambient conditions. However, examining Figure 7 leads to a substantially different conclusion. At the end, I wasn’t left with much confidence in the ability to predict organic phase, even for a fixed T, RH, and organic aerosol composition. The predicted ranges are exceptionally wide, even before considering uncertainty. I don’t have a specific suggested remedy, but I think it is something the authors need to address.

Address the applicability of using a parameterization based on pure compounds to predict properties of mixtures (e.g. SOA).

I second the editor’s comment regarding viscosity and T_g of complex mixtures. The authors should address, early in the manuscript, uncertainties in applying a parameterization based on pure compounds to SOA, which is a complex mixture. I don’t expect the authors to solve this, but it should be addressed early in the paper.

I also second the editors comment on Figure 7 and the discussion of T_g of PMF-derived factors. I’m not sure what it even means to have a predicted viscosity/ T_g of HOA, COA, MO-OOA, etc. since they are always mixed with other factors in the real atmosphere. Again, the authors can’t be expected to solve this, but more context in the text is needed early in the manuscript.

Address the uncertainty/error in the inputs to the parameterizations. I applaud the authors for making the best possible use of the available data and I understand they are trying to estimate properties of as many compounds as possible. However, I suggest they more thoroughly address errors in the data used to build the parameterizations (aside from applicability issues above) and

clarify when inputs are measured vs calculated. The authors provide prediction bounds, but my interpretation is that these are largely related to errors in fitting. What about errors in predicting T_g and Co ? How large are the errors associated with assigning a fixed factor of $T_g = 0.7 * T_m$ for a wide range of compounds? Can the authors compare measurements of T_g to T_g derived from T_m and show that plot in the SI? I browsed through the SI but didn't find reference to measured values of T_g , but surely there must be some measured T_g for pure compounds? How many compounds with both measured T_g and measured Co are in the training dataset? Are there any?

Can the authors provide a plot of measured Co vs estimated Co in the SI? Casual perusal suggests estimated Co can be very far from measured Co , which will in turn introduce error into the T_g prediction.

Lines 155-156 mentions that T_m is estimated, but most other references to T_m say it is measured (e.g. line 201). Is T_m calculated or it measured?

Lines 286-292. Lines 306-311 Are the volatility distributions derived only for the organic acid portion of the SOA? If so, how would this bias the measurements? Wouldn't organic acids be on the lower end of the volatility distribution, all else being equal?

How are the CIMS measurements biased or impacted by the ion chemistry employed? The reference indicates acetate ion chemistry was used and that it is sensitive to acids but not other SOA components. So, wouldn't the thermogram method be biased toward low volatility compounds?

The authors indicate that the formulas method is biased by decomposition of SOA. How does the thermogram method deal with decomposition products, which are indicated to be extensive?

Lines 331 and subsequent discussion on RH. It wasn't very clear how the T_g parameterization presented earlier in the paper relate to the calculations of RH effects. Is the T_g parameterization used in the calculation of T_g of the water/organic mixtures or are these calculations independent of each other? Can the authors show the equation for calculating T_g of water/organic mixtures, since this is central to the paper?

I found the transition between the bulk of the paper, which focuses on parameterizations of T_g as a function of volatility, to this section of the paper on the impact of RH somewhat abrupt. Can the authors comment on which effect (condensed phase water vs volatility) has a larger influence on organic phase state? For example, at line 321 the authors state that T_g was 313-330 K during SOAS, which would mean the aerosols are primarily in a glassy state. However, on lines 351-352 they state that the particles were mostly liquids.

Minor Comments and Technical Corrections

General comment. There are quite a few typos and grammatical errors through the manuscript. This doesn't get in the way of understanding the paper, but it was noticeable. I started to make

specific suggestions below but stopped after a few pages of text. A thorough editing would improve the paper.

Line 38. I think you mean SOA derived from diesel fuel rather than the viscosity of diesel fuel itself.

Line 62. Suggest changing “depending on” to “as a function of”. The message of this sentence is unclear, given you cite many measurements of particle phase via particle bounce.

Line 66. I think you mean “in the bulk organic phase” rather than “bulk organic molecules”.

Line 72. Add “The” before “Chemical”

Line 75. Add “a” before “phase”

Line 113-114. Isn't it more accurate to say that you parameterized the relationship between T_m and C_o ? As far as I can tell the vast majority of T_g values are estimated from measured T_m .

Line 123 add “or” after “measured”

Line 129. Can the authors state the fraction of the compounds with measured C_o and measured T_g as opposed to estimated values?

Line 132. The weak dependence of T_g on O:C is not very clear from Fig. 1a. Suggest graphing this separately, perhaps in the SI.

Lines 155-156. Are T_m values themselves estimate or taken from a database of measured compound melting points?

Line 252 change comparing to compared

Lines 289 – 295. It isn't clear how to reconcile 50% of the total OA being composed of organic acids (L289) if “many of the detected species are decomposition products (L294-295)”. Please clarify.

Lines 292-294. How did the authors convert from the molecular formulas measured by the MS to molecular structure needed for the group contribution methods?

Lines 387-389. I can imagine that the aerosol organic loading in Beijing is also significantly larger than at most other sites, which will impact volatility and T_g due to partitioning. Please comment.

Lines 459-461. The particles were solid-like when anthropogenic influence from the Manaus plume dominated.

Lines 474-477. This is true, but other studies measured or implied kinetic limitations at moderate or high RH. These studies should also be cited.

Figures 1-3. Isn't possible to differentiate among the symbol shapes.

Figure 1a. Can the fit from equations 1 and 2 be drawn on this figure?

Figure 3. Clarify whether Measured T_g is the same as T_g estimated from T_m as in Figures 1 and 2.

Figure 4 caption. It isn't clear what "edge lines are in grey" refers to.