Response to the comments of Anonymous Referee #1

Referee General Comment:

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.

Response: We thank Anonymous Referee #1 for the positive review and very helpful suggestions. Following your suggestions, we have improved the writing in the revised manuscript and added some statements to prevent an abrupt transition between different sections. We add a new subsection 2.1 "Dataset of glass transition temperature" to describe the training and test datasets and address the uncertainties and limitations in the inputs (see our response to your comment 3). We also divide the Method section into three subsections and keep the main parameterization predicting the glass transition temperature as a function of volatility in the main text. We move other parameterizations and related comparisons to the Appendix (see our response to your comment 1). In addition, we re-organize the paragraphs about the T_g of total OA at the dry condition ($T_{g,org}$) during the SOAS campaign and clarify that the most credible predicted $T_{g,org}$ values span in the range of 313 – 330 K (see our response to your comment 1). We also clarify the reasons why we predict the viscosity at different relative humidity and the relative importance of volatility and particle phase water in

OA phase state predictions (see our responses to your comments 1 and 5). We believe after addressing the above major issues, the take-home messages of this manuscript are clearer. Please see the detailed responses below.

Referee Major Specific Comment:
(1) 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.

Response: Following your suggestions, we have (1) added titles in Fig. 1, Figures in the Appendix and Figs. S3-5; (2) shortened the captions of all figures in the main text and the supplement; (3) referred to the parameterizations by descriptive names through the manuscript. In addition, we have moved the parameterizations as a function of elemental composition and the comparison with the parameterization in Zhang et al. (2019) to the Appendix for better readability of the main discussion.

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.

Response: In the revised manuscript we add a new subsection 2.1 "Dataset of glass transition temperature" to describe the training and test datasets and the methods deriving the values of T_g, C^0 and T_m. In the training dataset, T_g is measured or otherwise estimated from T_m. C^0 and T_m are estimated from the EPI Suite if they were not available from measurements. For the detailed description, please refer to the response to your comment 3. Figure 3 in the ACPD manuscript has been moved to the Appendix as Fig. B1. You are right that the data presented in this figure are a subset of the data in Fig. 1, which is clarified in the caption of revised Fig. B1. We add the following sentences in the revised manuscript:

Line 553-555: "Figure B1 compares the measured T_g included in the training dataset shown in Fig. 1a to T_g predicted by (a) C^0 and the atomic O:C (Eq. 1), (b) elemental composition (Eqs. A1-A3), and (c) Eq. (B1) by Zhang et al. (2019)".

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.

Response: This comment arises from the original Fig. 4 showing that the T_g of total OA under dry conditions is very different (span a range of 100 K as you pointed) predicted by different C^* distributions measured during the SOAS campaign. Note that this wide range stems mostly from variations in measured volatility distributions, but not from uncertainties of our T_g parameterizations and viscosity prediction method. In the ACPD manuscript we have stated that the credible $T_{g,org}$ values span in the range of 313 – 330 K and the predicted low values (< 280 K) estimated from the "Formulas" and "Partitioning" methods in Stark et al. (2017) are not credible. Our parameterizations can reasonably predict the T_g of ambient OA when measured C^* distributions are well constrained. The predicted viscosity of OA in SOAS is consistent with the ambient particle phase state measurements during the SOAS campaign (please see also our response for your comment 5). The T_g varying from 290 – 339 K stated in Line 31 in the abstract is the predicted T_g of total OA at the dry condition ($T_{g,org}$) at the eleven field sites. We clarified this point in the revised abstract. The pink shaded area in Fig. 7 (Fig. 5 in the revised manuscript) bounds the predicted viscosities of MO-OOA and LO-OOA in three different locations, thus the pink shaded area spans wide at medium / low RH, as the $T_{g,org}$ and hygroscopicity of these OA factors at different locations are different (please also refer to our response to your comment 2). We predict the viscosity of the OOA factor as OOA is often considered to represent SOA, and the predicted viscosity of OOA is consistent with the viscosity of SOA formed from various precursors. The above comparison shows the parameterization developed in this study can reasonably predict the phase state of ambient OA as well as laboratory-generated SOA. We add the following sentences in the revised manuscript:

Line 260-263: "Figure 2 shows that $T_{g,org}$ of total OA (TOA) range from 232 K to 334 K, depending on volatility distributions measured by different methods, while the most credible predicted $T_{g,org}$ values span in the range of 313 – 330 K. The reasons are stated below by comparing the different methods deriving the C^* distributions".

Line 303-308: "These analyses indicate that the volatility distributions derived from different methods, even when based on the same measurements, significantly affect the predicted $T_{g,org}$, and the most atmospherically relevant volatility distributions should be carefully chosen to reasonably predict the glass transition temperature of ambient OA. In summary, the $T_{g,org}$ values during the SOAS campaign should be in the range of 313 – 330 K".

(2) 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.

Response: Following your suggestion, we added a sentence to acknowledge additional uncertainty in the method section. We added a new subsection "2.3 Predictions of T_g and viscosity of organic aerosols" in the Method section. We assumed ideal thermodynamic mixing when applying the T_g parameterization to a certain volatility bin containing multiple components. After the T_g in a certain C^* is known, the T_g of SOA mixtures is calculated by the Gordon-Taylor equation and we state the limitation of the Gordon-Taylor equation in the revised manuscript. We add the following sentences:

Line 205 -208: "Note that there may be additional uncertainty in application of T_g parameterizations (which were developed based on pure compounds) to each volatility bin representing surrogate of complex multicomponent mixtures."

Line 228-233: "The Gordon-Taylor approach has been validated for a wide range of mixtures including SOA compounds (Dette et al., 2014; Lessmeier et al., 2018). The Gordon-Taylor approach may fail in the case of adduct or complex formation (Koop et al., 2011), which is highly unlikely in multicomponent mixtures with myriads of SOA compounds with very small individual mole fractions and thus particular interactions between individual compounds are more likely to average out (Shiraiwa et al., 2017); this aspect would need to be investigated in future studies".

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.

Response: We predicted T_g of OA factors as their volatility distributions were available and we think that comparison of T_g of these factors would be useful to compare with measured viscosities of laboratory-generated SOA, given that OOA may correspond to SOA (Jimenez et al., 2009). In the revised manuscript we have shortened the description of T_g of the characterized OA factors and explained the reasons why we compare the viscosity of OOA factors with the viscosity of laboratory-generated SOA. We also added a sentence to note potential limitations of this analysis.

Line 378-382: "Note that these different OA factors may often be internally mixed in ambient atmosphere and predicted $T_{g,org}$ and particle viscosity would be irrelevant in such a case. Nevertheless, these predictions can be useful when particles are externally mixed or ambient OA are dominated by a certain OA factor".

Line 406-407: "The predicted behavior of BBOA is in line with bounce measurements observing that particles are semisolid in a biomass burning plume (Bateman et al., 2017)".

Line 410-419: "There have been growing measurements of RH-dependent viscosity of laboratory-generated SOA formed from different precursors, e.g., isoprene (Song et al., 2015), α-pinene (Abramson et al., 2013; Renbaum-Wolff et al., 2013; Kidd et al., 2014; Pajunoja et al., 2014; Bateman et al., 2015; Zhang et al., 2015; Grayson et al., 2016; Petters et al., 2019), toluene (Song et al., 2016a) and diesel fuel (Song et al., 2019). As the OOA factors characterized from ambient AMS observations may represent ambient SOA (Jimenez et al., 2009), the predicted viscosities of OOA are compared with laboratory measurements of SOA viscosities in Fig. 5b. It shows that the majority of experimental values is well bounded by the predicted viscosities of OOA, represented by the pink shaded area".

(3) 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 C^o? 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 $\boldsymbol{C^o}$ are in the training dataset? Are there any? Can the authors provide a plot of measured $\boldsymbol{C^o}$ vs estimated $\boldsymbol{C^o}$ in the SI? Casual perusal suggests estimated $\boldsymbol{C^o}$ can be very far from measured $\boldsymbol{C^o}$, 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?

Response: In the revised manuscript we add a new subsection to describe the training and test datasets and the methods deriving the values of T_g, C^0 and T_m. The training dataset is used to develop the parameterizations and T_m is measured or otherwise estimated from the EPI Suite. The test dataset is used to validate the parameterizations and all of the T_m values are estimated from the EPI Suite. We have included Table S1 showing the number of compounds with their T_g, C^0 and T_m measured or otherwise estimated. Most CH, CHO, and CHON compounds have measured T_g, C^0 or T_m; relatively large uncertainty in the inputs data exists in CHOS compounds as only 1 CHOS compound has measured T_g (Zhang et al., 2019). To make our parameterizations also applicable to CHOS compounds, we include other CHOS compounds (Li et al.,

2016) with both C^0 and T_m estimated from the EPI Suite. We include Fig. S1 showing (a) the comparison of measured T_g and the T_g estimated by the Boyer–Kauzmann rule and (b) the comparison of C^0 measured and estimated from the EPI suite. We add the following sentences in the revised manuscript to address the uncertainty in our input data:

Line 127-153: "Measured T_g values are available for 42 CH compounds, 259 CHO compounds, 35 CHON compounds and 1 CHOS compound (Koop et al., 2011; Rothfuss and Petters, 2017; Lessmeier et al., 2018; Zhang et al., 2019), among which there are 168 compounds with measured C^0 available (Table S1). When T_g measurements are unavailable, T_g is estimated from the melting temperature (T_m) applying the Boyer-Kauzmann rule of $T_\mathrm{g} = g \cdot T_\mathrm{m}$ (Kauzmann, 1948; Boyer, 1954) with $g = 0.70085$ (± 0.00375) (Koop et al., 2011), referred to "estimated T_g" in this study (see good agreement of measured and estimated T_g in Fig. S1a). 1187 compounds (391 CH, 537 CHO, 241 CHON and 18 CHOS compounds) with both measured T_m and C^0 (Table S1, S2) are adopted from the MPBPWIN Program Test Sets (http://esc.syrres.com/interkow/EpiSuiteData.htm) included in the Estimation Programs Interface (EPI) Suite software version 4.1 (US EPA, 2015). Measured T_g, T_m or C^0 for CHOS compounds are sparse and we adopt 850 CHOS compounds included in Li et al. (2016) with their T_m and C^0 estimated by the EPI Suite software (Table S2). There are estimation limitations in the EPI Suite; for example, the disagreement between measured and estimated C^0 is larger for compounds with $C^0 < \sim 10^{-2}$ μg m^{-3} (Fig. S1b), which may affect the T_g predictions for compounds with low volatility. However, given the large amount of data points with measured C^0 included in the training dataset, the estimation bias introduced by the EPI Suite may not substantially impact the accuracy of the parameterization developed in this study.

The test dataset used to validate the performance of the parameterizations predicting T_g of SOA components includes 654 CHO compounds and 212 CHON compounds found in SOA oxidation products (Shiraiwa et al., 2014). The values of their C^0 are estimated using the EVAPORATION model (Compernolle et al., 2011). Their T_m values are adopted from the EPI Suite. The T_g predicted by our parameterizations are compared with the T_g estimated from the T_m applying the Boyer-Kauzmann rule in the test dataset".

Table S1. Number of the compounds included in the training dataset and their T_g, C^0 and T_m measured or otherwise estimated.

	CH	CHO	CHON	CHOS
Both T_g and C^0 measured	38	125	5	0
Measured T_g, C^0 estimated from EPI Suite	4	134	30	1
T_g estimated from measured T_m, measured C^0	391	537	241	18
T_g estimated from T_m, T_m estimated from EPI Suite, measured C^0	0	0	0	11
T_g estimated from measured T_m, C^0 estimated from EPI Suite	0	0	0	63
T_g estimated from T_m, T_m and C^0 estimated from	0	0	0	850

EPI Suite

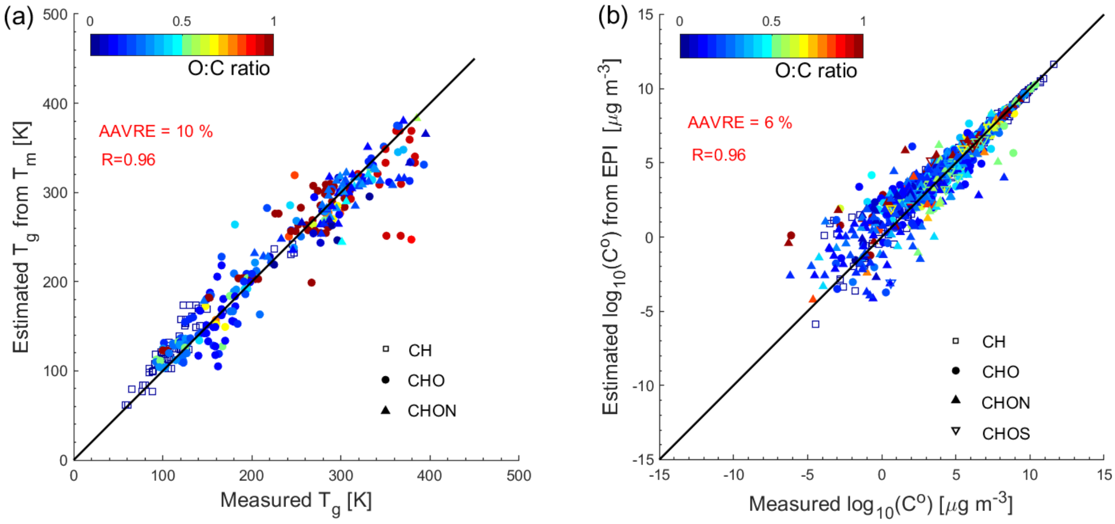

Figure S1. (a) Comparison of the measured T_g (Koop et al., 2011; Dette et al., 2014; Rothfuss and Petters, 2017; Lessmeier et al., 2018; Zhang et al., 2019) and the T_g estimated by the Boyer–Kauzmann rule for 336 organic compounds with their measured T_m available. (b) Comparison of pure compound saturation mass concentration (C^0) measured and estimated from the EPI suite for 1637 organic compounds included in the training dataset in this study.

(4) 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?

Response: It is correct that acetate CIMS mainly measures organic acids, besides a few other compounds such as phenols and other compounds with acidic hydrogen. While this indeed somewhat biases the measurements, the acid fraction in SOA has been shown to be high, as pointed out in the manuscript and referenced (Yatavelli et al., 2015). Further, we disagree that organic acids populate a particularly low volatility region since they show a broad range of O:C and carbon number range, similar to other SOA components such as alcohols, aldehydes, ketones, peroxides, etc. Lastly, the thermogram method certainly includes thermograms of species thermally decomposing rather than desorbing from the filters. However, the peak temperatures of these thermograms relate to volatilities closer to the species volatility than from the other methods and can therefore still be seen as better overall volatility measurements than the other two methods. We have added an extra sentence as below:

Line 272-275: “While this method may be influenced by thermal decomposition, the peak temperatures of decomposing species can be expected to relate closer to actual volatilities than any of the other two analysis methods (Stark et al., 2017)”.

(5) 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.

Response: The T_g parameterization is used in the calculation of T_g of the water/organic mixtures. In the revised manuscript we describe the procedures calculating the T_g of water/organic mixtures and the viscosity at different T and RH in the new subsection 2.3 “Predictions of T_g and viscosity of organic aerosols”. We added one sentence at the beginning of this paragraph to explain why we further predict the viscosity at given RH to prevent an abrupt transition between sections. The $T_{g,org}$ was predicted to be 313-330 K during the SOAS campaign, indicating that OA are primarily in a glassy state at the dry condition. To compare with the ambient phase state measurements, we calculate the viscosity at different RH and the average T during SOAS. The predicted viscosity is liquid at RH of 83 % and semi-solid at RH of 50 %, consistent with the particle bounce measurements. We add a few sentences in the revised manuscript discussing the influence of condensed phase water vs volatility on phase state.

Lines 235-249: “Under humid conditions, the water content in OA can be estimated using the effective hygroscopicity parameter (κ) (Petters and Kreidenweis, 2007). The T_g of organic-water mixtures ($T_g(w_{org})$) at given RH can be estimated using the Gordon-Taylor equation (Gordon and Taylor, 1952):

$$T_g(w_{org}) = \frac{(1-w_{org})T_{g,w}+\frac{1}{k_{GT}}w_{org}T_{g,org}}{(1-w_{org})+\frac{1}{k_{GT}}w_{org}} \qquad (6)$$

where w_{org} is the mass fraction of organics in particles; $T_{g,w}$ is the glass transition temperature of pure water (136 K, Kohl et al., 2005), and k_{GT} is the Gordon-Taylor constant for organic-water mixtures which is suggested to be 2.5 (Zobrist et al., 2008; Koop et al., 2011). Viscosity can then be calculated applying the Vogel-Tammann-Fulcher (VTF) equation (Angell, 1991): $\eta = \eta_\infty e^{\frac{T_0 D}{T-T_0}}$, where η_∞ is the viscosity at infinite temperature (10^{-5} Pa s, Angell, 1991), D is the fragility parameter which is assumed to be 10 (DeRieux and Li et al., 2018), and T_0 is the Vogel temperature calculated as $T_0 = \frac{39.17\,T_g}{D+39.17}$”.

Line 318-320: "We further calculate the viscosity of OA based on the $T_{g,org}$ of TOA predicted above in order to compare with the ambient phase state measurements during the SOAS campaign".

Line 337-342: "The variations (313 – 330 K) in $T_{g,org}$ due to the different measured C^* distributions (Fig. 2) have a more significant impact on the predicted viscosity at low and medium RH (Fig. 3a). When RH is higher than ~70 %, the predicted viscosities calculated from different $T_{g,org}$ values are very close; at high RH the condensed phase water has a larger influence on the phase state than the volatility does, depending on the hygroscopicity of organic aerosols".

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.

Response: Thanks for reading our manuscript carefully and the specific suggestions. We have done a thorough editing.

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

Response: Right. The "of" in front of diesel fuel has been deleted.

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.

Response: "depending on temperature (T), relative humidity (RH), and chemical composition" has been moved to Line 60.

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

Response: The sentence has been changed to "bulk diffusivity of organic molecules".

Line 72. Add "The" before "Chemical"

Response: Added.

Line 75. Add "a" before "phase"

Response: Added.

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.

Response: The object of this study is to develop parameterizations predicting T_g and further predict the viscosity of ambient organic aerosols. T_g values estimated from T_m are used as part of our training dataset to develop the parameterization for T_g as a function of C^o. The parameterizations perform well predicting T_g of individual organic compounds (Fig. 1), the viscosity of ambient OA (Fig. 3) and laboratory generated OA (Fig. 5), indicating that estimated T_g values from T_m used in the training dataset would not impact predictions of T_g by the parameterizations developed in this study. Please also refer to our response to your comment 3.

Line 123 add “or” after “measured”

Response: Added.

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

Response: The fractions of the compounds with measured C^o and measured T_g as opposed to estimated values in the training dataset are different for different classes (CH, CHO, CHON and CHOS). We have included Table S1 showing the number of compounds with their T_g, C^0 and T_m measured or otherwise estimated. Please see our response to your comment 3.

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.

Response: The dependence of T_g on O:C is included in the SI in the revised manuscript:

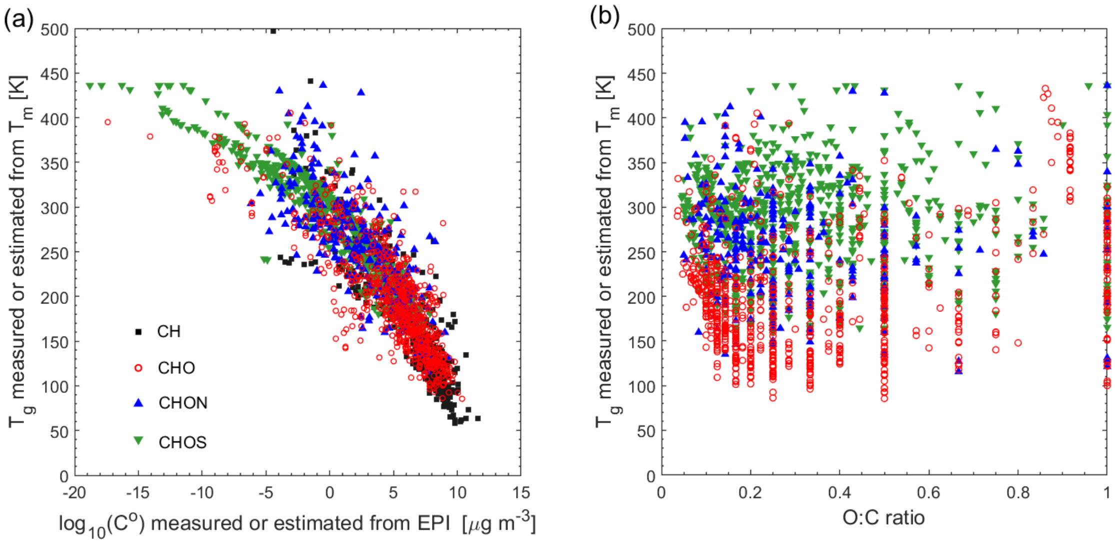

Figure S2. T_g of organic compounds in the training dataset plotted against (a) pure compound saturation mass concentration (C^0) and (b) the atomic O:C ratio.

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

Response: There are two datasets used in this study. The training dataset is used to develop the parameterizations and T_m is measured or otherwise estimated from the EPI Suite. The test dataset is used to validate the parameterizations and T_m values are all estimated from the EPI Suite. Please also refer to our response to your comment 3.

Line 252 change comparing to compared

Response: Changed.

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.

Response: The FIGAERO measurement are not used to determine the overall OA composition or, for that matter, the contribution of organic acids to total OA. The 50% value was taken from a separate study, where acetate CIMS measurements were compared quantitatively to AMS measurements (Yatavelli et al., 2015).

We clarified at Line 265: …organic acids (which were shown to account for about half of the total OA; Yatavelli et al., 2015) from…

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?

Responses: The details of this algorithm are described in Stark et al. (2017). Briefly, at least one carboxylic acid group was assumed to be present in each formula, while the remaining oxygen was either assumed to be carbonyl, hydroxyl, or carboxyl. The resulting volatility distributions from these three different possibilities were considered in the Stark et al. (2017) paper, but all showed similar distributions in that the volatilities were all very high. In this study, we used the results from the assumption that all remaining oxygen was present in hydroxyl groups, resulting in the lowest volatilities of the possible formula method results. We added a sentence describing a few more details:

Line 272-275: "While this method may be influenced by thermal decomposition, the peak temperatures of decomposing species can be expected to relate closer to actual volatilities than any of the other two analysis methods (Stark et al., 2017)".

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.

Responses: Following your suggestion, we add the following sentence:

Line 385-387: "This may be due to the higher total OA mass concentrations in Beijing (Xu et al., 2019), which facilitates greater partitioning of SVOC compounds into the particle phase, leading to a lower $T_{g,org}$".

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

Response: We revised this sentence at Line 460:

"……while with the anthropogenic influence including both urban pollution and biomass burning, they occur as semi-solid or glassy (Bateman et al., 2016; Bateman et al., 2017)".

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.

Response: This sentence has been revised as:

Line 485-489: "Some chamber experiments probing the mixing timescales of SOA particles formed from isoprene, *α*-pinene, and limonene did not observe significant kinetic limitations at moderate and high RH under room temperature (Loza et al., 2013; Ye et al., 2016), while kinetic limitations of bulk diffusion of organic molecules in *β*-caryophyllene SOA have been observed at 75 % RH (Ye et al., 2018), warranting further investigations on the degree of kinetic limitations in ambient tropospheric conditions".

Figures 1-3. Isn't possible to differentiate among the symbol shapes.
Response: The size of the symbols has been enlarged and the resolution of the figures has been improved.

Figure 1a. Can the fit from equations 1 and 2 be drawn on this figure?
Response: The fit from Eq. 1 and 2 has been added in Fig.1a and Fig. S3a, respectively. We add the following sentences in the Line 171-174:
"The predicted T_g by Eq. (1) is plotted in Fig. 1a with the O:C ratios of 0, 0.5, and 1, showing that the predicted dependence of T_g on C^0 follows the trend well in the training dataset. The O:C ratio mainly affects the predicted T_g of volatile or extremely low volatile compounds".

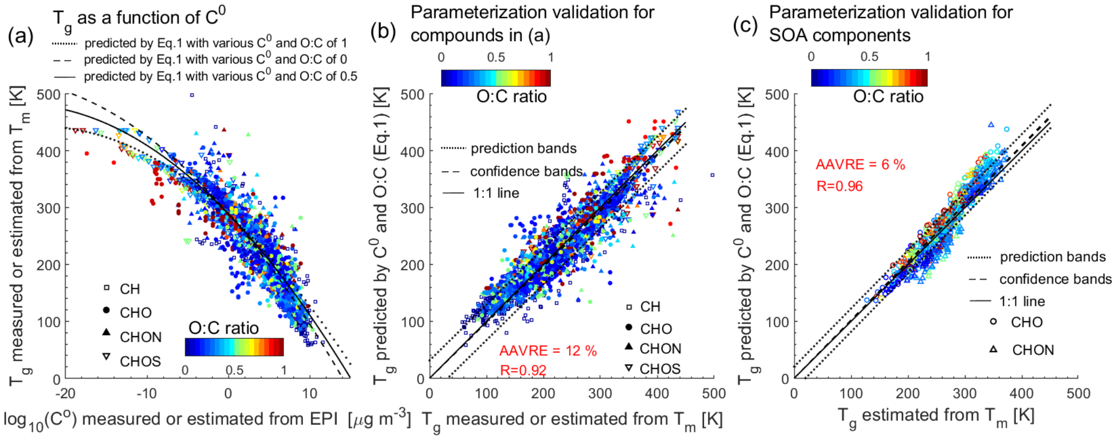

Figure 1. (a) T_g of organic compounds in the training dataset plotted against C^0. The lines show the predictions of T_g (Eq. 1) by C^0 and the O:C ratio of 0 (dashed), 0.5 (solid), and 1 (dotted). (b) Predicted T_g by C^0 and the O:C ratio (Eq. 1) for compounds shown in (a) compared to measured or otherwise estimated T_g from T_m. (c) Predicted T_g for SOA components (Shiraiwa et al., 2014) using Eq. (1) plotted against estimated T_g from T_m with the Boyer-Kauzmann rule. The correlation coefficient (R) and the average absolute value of the relative error (AAVRE) are shown. The dashed and dotted lines in (b) and (c) show 68% confidence and prediction bands, respectively.

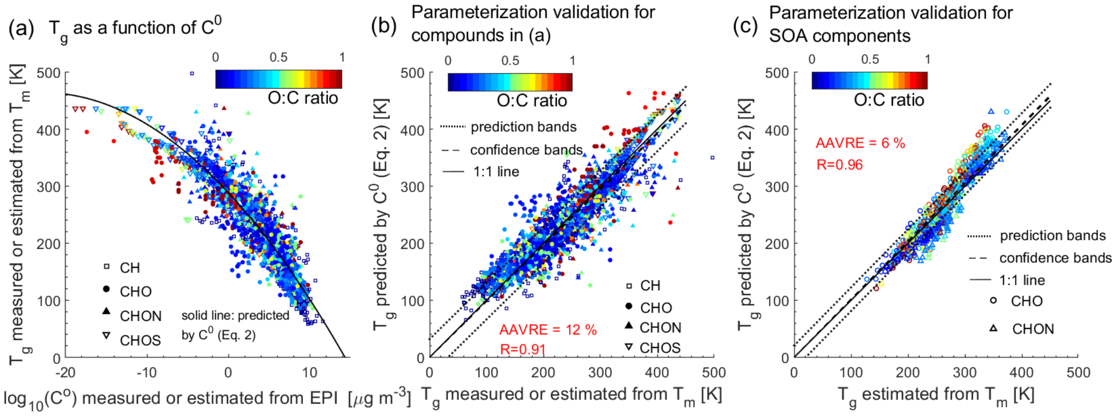

Figure S3. (a) T_g of organic compounds in the training dataset plotted against C^0. The solid line shows the predictions of T_g by C^0 (Eq. 2). (b) Predicted T_g by C^0 (Eq. 2) for compounds shown in (a) compared to measured or otherwise estimated T_g from T_m. (c) Predicted T_g for SOA components (Shiraiwa et al., 2014) using Eq. (2) plotted against estimated T_g from T_m with the Boyer-Kauzmann rule. The correlation coefficient (R) and the average absolute value of the relative error (AAVRE) are shown. The dashed and dotted lines in (b) and (c) show 68% confidence and prediction bands, respectively.

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

Response: The comparison between the measured T_g and the T_g estimated from T_m has been shown in Fig. S1. Please refer to our response to your comment 3. The original Figure 3 has been moved to the Appendix B and revised clarifying the measured T_g are the same as the measured T_g shown in Fig. 1a.

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

Response: It has been changed to "marker edge lines are in grey".