

## ***Interactive comment on “A two-dimensional volatility basis set: 1. organic-aerosol mixing thermodynamics” by N. M. Donahue et al.***

**N. M. Donahue et al.**

nmd@andrew.cmu.edu

Received and published: 9 February 2011

We have substantially reorganized the manuscript in response to comments from Reviewer 2, so many equation and figure numbers in the revised manuscript have changed.

*Review of Donahue et al. “A two-dimensional volatility basis set: 1. Organic-aerosol mixing thermodynamics”*

*Donahue et al. present a very interesting and detailed description of a “2-D basis set” of organic aerosol that relates chemical composition (specifically, O:C) to compound volatility and activity. As the authors are no doubt aware, their manuscript is extremely dense, but they do a very good job of explaining the meaning of the equations they*

C13379

*present. Given the “ease” with which O:C measurements are now made, this paper should have a substantial impact on the field. I ask the authors to consider the following comments prior to publication. Equation 8: Based on the definition of  $\delta\theta_{i,s}$  given in Eqn. 4, if I plug this into Eqn. 8 I come up with*

$$\theta_{i,s}^{vap} = \theta_{i,i} + 2\delta\theta_{i,s} = \theta_{i,i} + 2 * (2\theta_{i,s} - \theta_{s,s} - \theta_{i,i}) = 4\theta_{i,i} - 2\theta_{s,s} - \theta_{i,i}$$

*This is different than what is given in Eqn. 3:*

$$\theta_{i,s}^{vap} = 2\theta_{i,s} - \theta_{s,s}$$

*Am I missing something here?*

Eq. 4 had a typographical error. It should have read:

$$2\delta\theta_{i,s} = 2\theta_{i,s} - \theta_{s,s} - \theta_{i,i}$$

in which case the substitution is straightforward. We now simply define the average  $\bar{\theta}$  and deviation  $\delta\theta_{i,s}$  and proceed. This has all been moved to an appendix.

*P. 24100, L. 4: It is mentioned that the total excess energy varies quadratically with the fraction of “A” in the solution. I think this result comes from the use of a simple two component mixture. For a multi-component mixture, the quadratic relationship would be changed. Assuming I’m right, I think this should be mentioned. Of course, the main development of the manuscript is essentially a two-component model (oxygen and carbons), and thus the two-component treatment here is most relevant to the later sections.*

The subsection this development sits in is now entitled “Simple two-component mixtures”.

*P. 24102, L. 2: It is stated that for now the authors are neglecting “other” contributions beyond oxygen and carbons. In addition to mentioning N and S atoms here, I think*

C13380

*it would be useful to mention that the model also neglects to some extent functional groups. More specifically, the authors assume that OA will always have some distribution of functional groups that will lead to a specific value for the VP decrease per oxygen added. Although this is developed from a treatment where different functional groups are considered, the end result is that differences in functional groups are “neglected” in the model because oxygen addition is treated in an average sense.*

We attempted to make this point quite clearly in the original manuscript, but we have strengthened it in the revised manuscript. We are relying on evidence from numerous sources that the ratio of -OH and =O functional groups is roughly 1:1 to collapse the effect of multiple oxygenated functional groups into a single average value for O. Because we now lead with the empirical group contribution method, this should now be even more evident.

That being said, this is a very, very important point. We are simply not trying to present this framework for organic compounds in general. Counterexamples are easy to find, but it is not at all obvious that those counterexamples are atmospherically relevant. Even individual molecules that do appear in atmospheric organic aerosol will probably deviate from the average behavior we seek to describe, because the average is exactly that – an average. Instead, we are trying to answer a question: “Can we reasonably constrain organic-aerosol composition in the atmosphere based on only two measurable properties, volatility and the extent of oxygenation?” This question is now stated in the abstract.

*P. 24103, L10: With regards to the deviation of the large alkanes from the trend line determined from the smaller alkanes, the authors suggest that they believe that the difference is from measurement error and that, in fact, their extrapolation is more accurate. To me, this seems too bold of a statement. There are at least some reasons to think that as molecules become bigger (and more “floppy”) the ability of linear SAR’s to predict vapor pressures might start to fail to some extent. Nonetheless, given the estimated range of carbon numbers for the compounds that comprise “OOA” (6-14),*

C13381

*inaccuracies for alkanes with  $n_c > 20$  probably will have minimal impact on outcomes using the 2-D model.*

This is a fair point, and we have tempered our language in this section. However, we have also added some examples of atmospherically relevant compounds that fall in the part of the 2D space we previously described as based entirely on extrapolation. Specifically, we point out that the sub-cooled liquid saturation concentration of levoglucosan appears to be within about 1 order of magnitude of the average for C<sub>6</sub>O<sub>5</sub> compounds given by our analysis.

However, there is some reason to want a reasonable representation for the larger carbon number hydrocarbons, as they probably constitute a fair fraction of the HOA observed by the AMS, or primary organic aerosol in general. Our argument in this paper is that the activity coefficients of these large, nonpolar molecules in the background, much more polar oxidized organic aerosol (OOA) is important because high activity in the mixtures can keep the semi-volatile parts of the HOA fraction in an external mixture on distinct particles (for example a traffic mode of POA in sub-100 nm particles).

We argue that the question of whether POA particles are good seeds for SOA condensation is largely a red herring, based on the outdated notion that POA consists of mostly non-volatile organics. As it is now evident that most POA constituents are relatively volatile (and also that through much of the atmosphere POA comprises much less than 50% of the total OA), the relevant question is really whether POA constituents dissolve well in OOA – whether SOA is a good seed for POA.

*Eqn. 16: It would be useful to remind the readers that  $n_C^0 = 25$  is by definition.*

This is now Eq. 2. It appears at the end of a discussion detailing how each of these three parameters is estimated, so as currently written we believe that the origin of all three parameters is clear.

*Diacids: The authors use the diacid VP’s from Cappa et al. (2007). (Presumably these*

C13382

are the sub-cooled values and not the solid-state values.) However, there is a fair amount of debate in the literature as to the “true” vapor pressures of these compounds and the Cappa et al. results have tended to give the lowest estimates, at least for diacids with 7 carbons or greater (and with the exception of the single data point from Yatvelli and Thornton (2010)). This may be a reason for the difference between the measurements and the SAR results (dashed lines in Figure 3). Whether or not this is the case, it seems that some brief mention of the “controversy” over diacid vapor pressures is in order and, perhaps, some justification for this choice. For example, the authors might point out that the slope of  $\ln(VP)$  vs  $n_C$  for the Cappa et al. dataset is generally consistent with expectations based on the trends for the alkanes, while for some of the other datasets the slope is somewhat shallower, tending towards zero in some cases. I think that this could have some influence on the determination of the exact magnitude of the  $\delta\theta_{CO}$  term, as this seems to come from an empirically determined value for  $b_{CO}$  (which was taken as -0.3). However, the measurements for the diacids do not show as much variability at low carbon number ( $n_C < 7$ ), and this is where the curvature in the non-linear SAR really starts to come into effect. Therefore, the determination of the non-linear terms may not be dramatically influenced by the uncertainties in the diacid vapor pressure measurements because such uncertainties are smaller for small diacids. Ultimately, depending on whether the model is being used in a diagnostic or prognostic sense, this will lead to an uncertainty in deduced parameters such the mean MW (or molecular volume) or in the predicted OA mass.

We do use the Koponen and Cappa values. As suggested we have added some discussion of the lingering uncertainty in the diacid vapor-pressure measurements. Because of the way we formulate it, the  $b_{CO}$  term does influence both the low and high  $n_C$  values for the diacids, but the reviewer is correct that the influence is greatest at low  $n_C$ , where the disagreement among the various measurements is least. Looking at the non-linear group-contribution values in Fig. 3, our prediction for larger  $n_C$  does fall roughly a factor of 3 above most of the plotted points. We have added discussion as suggested in three places – where the diacids are first introduced in the middle of

C13383

section 2.1, just before subsection 2.1.1 when the non-linear term is being introduced, and finally again at the end of subsection 2.1.1.

*Figure 3/7: What is going on with the monocarboxylic acids?*

The NIST data show a lot of scatter above C8.

*Figure 5: I find the boxes with the numbers (e.g. 6, 4, 2) to be a little confusing. They appear to correspond to the green lines, which indicate the number of oxygen atoms. However, because the green, cyan and black curves all intersect at 1:1, it makes it appear as if the cyan curves are labeled with the same number as the green (or black) curves. However, I think that in reality at 1:1 the cyan curves correspond to twice the number in the box (because  $n_O = n_C$  at this point). So then, from right to left on the figure, the cyan correspond to 4 atoms, 6 atoms, 8 atoms, etc. Can this be clarified?*

Point well taken. We completely removed the  $n_M$  curves from the plot and added labels to the  $n_O$  isopleths as well.

*Figure 5: The authors have for the lower limit on the x-axis  $\log_{10}(C_0) = -5$  and draw their dashed box down to this region (with an upper limit for the dashed box at 1). I assume there a basis for choosing -5 as the lower limit for the dashed box, but it does not seem to be stated. This is only important because it effectively helps define the boundaries of the  $n_C/n_O$  pairs that are possible to give a particular O:C. If the box was cut at a less negative value (say, -2), then we would not expect to find compounds with e.g.  $n_C = 8$  and  $n_O = 6$ . But if it were allowed to extend down to lower values then the range of possible compounds would increase. Effectively, the boundaries for the box define the range of possible molecular weights for the OOA compounds and, although O:C is constrained from measurement, it is not clear how the x-axis values are constrained (if they are).*

There is abundant evidence that OOA extends to at least -5, but we really should allow that box to extend off the frame, as the lower bound is uncertain. This is explored in

C13384

depth in Cappa and Jimenez 2010, as we point out in the discussion. However, we redrafted the figure with the OOA box extending off-scale to the left.

*P. 24107, L. 6: The authors mention they are trying to define interaction energies between molecule and solvent functional groups. However, I think functional groups might be the wrong word in the context of the model, since information as to the specific nature of the functional groups has been consolidated into a single parameter. I think the authors are actually defining interaction energies between oxygens and carbons (without specific consideration of functional group) in the solvent and molecule. Similarly for P. 24108, L. 3: I think here the actual assumption is that the solute?solvent interactions are dominated by the relative O and C concentrations (as opposed to functionalities), and not the specific molecules.*

We believe that either statement is accurate. The key finding is that the relative abundance of -OH and =O functional groups appears to be roughly constant (near 1:1) for most OA mixtures. We are assuming that the contributions of these two functional groups can be averaged (and that the contribution of other functional groups like -OOH and -O- will average in without too much error).

*Eqn. 19, etc.: Some of the terms (e.g.  $\theta_{i,i}$ ) used in various equations seem to have slightly different meaning. For example, in 19 the solvation energy is for the linear SAR, whereas in Eqn. 22 it is for a non-linear SAR. Perhaps it would be helpful to somehow indicate that one is for the non-linear and one for the linear SAR (for example,  $\theta_{i,i}^{NL}$  vs  $\theta_{i,i}^L$  or something like that). This would help to keep the reader from wanting to compare things that are not directly comparable. Then (I think), one could lead the reader through a bit better, for example by giving*

$$\theta_{i,i}^{NL} = \theta_{i,i}^L + 2 \frac{n_C n_O}{n_C + n_O} \delta\theta_{CO}$$

*assuming, that is, that I've indeed understood this correctly. This way the cross term looks like a perturbation to the linear SAR.*

C13385

We hope that the revised format of the paper has made this clearer. When one formulates the problem with interaction terms, it is actually difficult to get a linear expression (i.e. SIMPOL) because the cross term emerges naturally. This is why we argue that the linear expression is a simplification of the more natural non-linear one rather than the non-linear expression being a perturbation about the linear expression. Thus, the (new) Eq. 6 and 9 ARE the same terms, but 6 only holds when 7 is true – when the cross term is exactly the average of the diagonal terms.

*P. 24109 L. 16: The authors state the carbon and oxygen moieties have “very similar” masses. I feel this is a bit of an overstatement, as the relative difference between 16 and 12 is actually pretty large ( $16/12 = 1.33$ ). So the use of “very” might be a bit of a stretch. The point of the section is nonetheless well taken.*

An the point of the comment is well taken. We toned down the language.

*P. 24111, L. 13: Here, the authors use  $\Delta\theta_{10} = 690$  K. According to page 24097, this value comes from “simple theoretical considerations”. However, as mentioned, empirical correlations give a larger value of  $\sim 1500$  K. Given that the entire exercise in this manuscript is based on empirical correlations between carbon/oxygen number and vapor pressure, it would seem only right to use the empirical value. If this is true, then  $\theta_{CC} = 713$  K and  $\theta_{OO} + 2\delta\theta_{CO} = 2550$  K. The  $b_{CO}$  term would consequently be affected. If the authors believe that the “theoretical” value is preferable over the empirical value, then this must be rigorously justified, because as it is written now it seemed like it was being argued at the beginning that the empirical value was preferable (e.g. it was assumed “precise”).*

The precision we were referring to is that  $\log_{10} C^o$  and  $\Delta H^{vap}$  are highly correlated. The slope of that relation is  $\Delta\theta_{10}$ . We have revised the presentation of activity coefficients (see below) to make the role of  $\Delta\theta_{10}$  explicit. As we state in the text, the qualitative features of Fig. 6 (new – old Fig 9) do not change dramatically even with a factor of 2 change in  $b_{CO}$ . Also, note that the activity coefficient at 300K depends on  $b_{CO}$  only, as

C13386

we now show. However, this issue is quite relevant to the change in activity coefficients with temperature.

*Figure 8: It seems as if at  $n_C = n_O = 4$ ,  $lC^0$  is 3, not 2 as stated in the caption and the text.*

This is because the figure was switched, as we pointed out in our comment. We now show only the non-linear group-contribution figure, as Fig. 4, and for  $n_C = n_O = 4$ ,  $\log_{10} C^0$  is 2.

*Section 4.0: I am having a very difficult time understanding how the activity coefficient calculations have been done. I started by trying to calculate backwards: from Eqn. 8 one can determine the value of  $\delta\theta_{i,s}$  that corresponds to a given activity coefficient. For example, for  $\gamma_{i,s} = 2$ ,  $\delta\theta_{i,s} = -103$  K. Then, given that for the example where O:C = 0.75:1 so  $f_c^s = 1/1.75 = 0.57$  and  $\delta\theta_{CO} = -207$  K, one can in principle solve for the value of  $f_c^i$  that will give  $\delta\theta_{i,s} = -103$  K for a given value of  $n_M$  using equation 28. However, when I try to do this, I find that for all  $n_M$  except  $n_M = 2$ , I cannot find a satisfactory solution, i.e. there is no value of  $f_c^i$  for which  $\delta\theta_{i,s} = -207$  K. Given these difficulties, it would be very useful if the authors were to provide some sort of sample calculation that shows explicitly how both the activity coefficient and the x-axis ( $\log_{10} C^0$ ) values were determined. This would help for the reader to follow. The reason I started this exercise was to understand the origin of the curvature in the activity coefficient contours in Fig. 9. It is not overly clear where this curvature arises from.*

The forward calculation is a lot easier. The advise is very well taken and we now make the substitutions to derive the 2D-VBS activity coefficient expression as follows:

$$\log_{10} \gamma_{300} = -\frac{2 \delta\theta_{i,s}}{\Delta\theta_{10}}$$

and for the organic aerosol case we have Eq. 11:

$$\delta\theta_{i,s} = n_M \left[ (f_c^i)^2 + (f_c^s)^2 - 2 f_c^i f_c^s \right] \delta\theta_{CO}$$

C13387

which combine to give

$$\log_{10} \gamma_{300} = -2 n_M \left[ (f_c^i)^2 + (f_c^s)^2 - 2 f_c^i f_c^s \right] b_{CO}$$

so for  $b_{CO} = -0.3$

$$\log_{10} \gamma_{300} = 0.6 n_M \left[ (f_c^i)^2 + (f_c^s)^2 - 2 f_c^i f_c^s \right]$$

so, for the case the reviewer selects,  $f_c^s = 0.57$ , let's start with an example compound  $\log_{10} C^o = -5$ , O:C = 0.20 ( $f_c^i = 0.833$ ), where  $n_C = 20, n_O = 4, n_M = 24$ . Plugging in,

$$\log_{10} \gamma_{300} = 0.6 \cdot 24 \cdot (0.6944 + 0.3249 - 0.95) = 0.6 \cdot 24 \cdot 0.069 = 1.008$$

Or  $\gamma = 10$ . This is approximately where the green contours in Fig. 6 cease, as the activity coefficient is 10. We include this example calculation in the revised manuscript.

The curvature arises from the  $n_M$ , which decreases rapidly as one moves to the right in these figures but also decreases, more gradually, as one moves upward in y. This is why we supplied the contours of  $n_M$  in the original 2D figures.

*Parting Thought #1: After reading through the paper a few times, I do feel as if the overall discussion could be enhanced through some case?study type examples. Although the authors do a good job of describing the 2?D VBS framework, in the end I found it still somewhat difficult to understand exactly how this type of methodology will be implemented in either a prognostic or diagnostic framework in a way that will ultimately decrease uncertainties with respect to OA formation and evolution in the atmosphere. However, the authors do mention that they are working on separate manuscripts that provide the type of case study that I would ideally like to see and I will await this future work.*

The case studies are the next paper, which will be posted soon. It just takes too long to work through the examples in the same manuscript as this; however, "Parting Thought #2" is a perfect example of exactly this.

*Parting Thought #2: The authors have been big proponents of thinking that SOA formation from largeish "intermediate volatility" hydrocarbons (from evaporated POA). However, these compounds would have a number of carbon atoms that seems large compared to the OOA spaced as constrained by observations. This would seem to suggest a very significant role for fragmentation in any chemical mechanism that would occur for only "lightly" oxygenated organics. It would be interesting to hear the authors thoughts on how efficiently, e.g., a C18 hydrocarbon (with  $\log C^o = 3.5$ ) would fragment into a, e.g., C8 hydrocarbon with only 4-6 oxygens per C8 hydrocarbon. If the C18 hydrocarbon had 8-12 oxygens (and then split right down the middle into two C9 hydrocarbons with 4-6 oxygens each), the oxygenated C18 molecule would have had a vapor pressure so low that it would be 100% in the particle phase and thus unable to go through gas-phase oxidation mechanisms to produce OOA. Does this suggest an important role for heterogeneous chemistry? I suppose another way to ask the same thing is to point out that if that C18 molecule had only 4 oxygens, the VP would be so low ( $\log_{10} C^o = -4$ ) that it would be entirely in the particle phase. So how and when does fragmentation occur for this long-chain lightly oxygenated hydrocarbon? As with the previous comment, I suppose that this question might best be addressed in future work where the authors use the model (as opposed to here, where they have developed the model framework). But I would be curious to know what they think.*

This is an excellent example of using the 2D space

---

Interactive comment on Atmos. Chem. Phys. Discuss., 10, 24091, 2010.

C13389