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 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

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

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?

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 (oxygens and carbons), and thus the two-component treatment here is most relevant to the later sections.

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

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

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

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

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

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?

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).

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

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, $\theta_{i,i}^{L}\ vs.\ \theta_{i,i}^{NL}$ 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

$$\theta_{ii}^{NL} = \theta_{ii}^{L} + 2\frac{n_c n_o}{n_c + n_o}\delta\theta_{CO}$$

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

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.

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 ~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 θ_{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").

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.

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, $\delta\theta_{l,s}$ = -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.

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.

Parting Thought #2: The authors have been big proponents of thinking that SOA formation from large-ish "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^0 = 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 C^0 = -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.

Chattopadhyay, S., and Ziemann, P. J.: Vapor pressures of substituted and unsubstituted monocarboxylic and dicarboxylic acids measured using an improved thermal desorption particle beam mass spectrometry method, Aerosol Science and Technology, 39, 1085-1100, 2005.

Bilde, M., Svenningsson, B., Monster, J., and Rosenorn, T.: Even-odd alternation of evaporation rates and vapor pressures of C3-C9 dicarboxylic acid aerosols, Environmental Science & Technology, 37, 1371-1378, 2003.

Salo, K., Jonsson, A. M., Andersson, P. U., and Hallquist, M.: Aerosol Volatility and Enthalpy of Sublimation of Carboxylic Acids, Journal of Physical Chemistry A, 114, 4586-4594, 10.1021/jp910105h, 2010.

Saleh, R., Walker, J., and Khlystov, A.: Determination of saturation pressure and enthalpy of vaporization of semi-volatile aerosols: The integrated volume method, J Aerosol Sci, 39, 876-887, 10.1016/j.jaerosci.2008.06.004, 2008.