Atmos. Chem. Phys. Discuss., 9, C6170–C6174, 2009
www.atmos-chem-phys-discuss.net/9/C6170/2009/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "The critical assessment of vapour pressure estimation methods for use in modelling the formation of atmospheric organic aerosol" *by* M. H. Barley and G. McFiggans

Anonymous Referee #2

Received and published: 25 October 2009

My overall assessment for this paper is minor revisions. I was quite impressed with the authors' work. I particularly enjoyed the care and rigour of their methodology in creating datasets for testing vapour pressure methods, and their application of these methods to Donahue's partitioning to show the potential for variations in vapour pressure calculation to impact organic aerosol formation. I also found the authors' approach of attempting to simplify a group method (rather than make it more complex) very interesting, along with the discussion on overfitting. I have some comments on this last issue below that I would like to see addressed in the final version of the manuscript, and a few points of clarification needed in their methodology and analysis discussion.

The issue of overfitting, introduction: Introduced at page 18380, lines 5 to 10, and discussed briefly lines 22 to 26. I think the paper would benefit from a more thorough discussion of the underlying reasons why overfitting might be an issue for the creation of group methods for vapour pressure datasets for multifunctional compounds. I can think of a few possibilities: (a) vapour pressure and/or boiling point measurement errors may be large for any individual compound, hence group methods based on these measurements will be in error; (b) group-group interactions may sometimes have a large effect on the vapour pressure. The first of these would argue for an approach other than that taken by the authors in terms of the number of compounds used to create a group method. The authors argue for a smaller number in order to avoid overfitting, but with a smaller number of samples, the potential for the errors associated with any one sample becomes large. Thus if only two samples are used to describe the contribution of a specific group, and large measurement errors occur for one sample, the resulting vapour pressure estimates for that group may be large. Possibility (b) implies that methods that do not account for the interactions will be inaccurate and/or compounds with interacting groups would not be best for use in constructing the method. The paper would be improved with the addition of a paragraph or two after line 4 on page 18379 with a more detailed discussion of the "overfitting" problem, with some examples. The authors should attempt to describe the problem in a general mathematical sense as well; how does using all of the available information lead to poor predictions? The concept is counter-intuitive, and needs to be better explained. On a similar vein: the authors create a simplified version of Nannoolai et al's group method without really discussing their reasons for the simplification or why they expect it to perform better than the original. This reasoning would greatly improve the paper.

Methodology comments: Page 18381, lines 15 to 20. The authors dismiss a group of methods without adequate justification at this point in the discussion. I think they demonstrate the issue later in the paper, though its not clear whether the methods of Poling et al 2001 are similar to the ones compared later in the current work. If not, how can the authors be sure of this statement, not having checked out these methods against observations? If so, reference should be made to that demonstration here ("As we demonstrate later in this work…") rather than just say that the methods are "unlikely" to work well. If a class of methods have been excluded a priori from the comparison that occurs later, the authors should take care to note that in their abstract and conclusions, and temper or caveat their findings accordingly.

Page 18384, equation (9): the equations should be modified so that the same variable is used for the same quantity throughout all the equations. The Trb introduced here is apparently the same as the Tp introduced earlier.

Page 18385, lines 1 through 10: The issue of simplification of more complex methods and the reasoning behind the choices made needs to be more adequately explained. Table 1 summarizes what was done, but the text does not give a description or justification for what was done. Why not just use the original groups?

Page 18387. Lines 17 – 26. I appreciate the authors' approach to ensuring that the test data are of adequate accuracy for comparison to the different methods. They should include a discussion of which species / species classes / species groups are excluded from the test dataset by their approach.

Page 18388, top paragraph. The selection of no more than two samples from a functional group in generating the test set strikes me as being a problem: if one or both of these measurements has significant errors, will this not result in large apparent biases of the predictions? This is a counter argument to concerns over overfitting that the authors should address.

Page 18388, line 16, condition number 1. This is a potentially serious omission. Both alkenes and aromatics have significant yields of secondary organic aerosols from the oxidation of these precursors. The oxidation products may include unbroken ring groups and double bonds. Their exclusion here implies that large classes of compounds known to be significant organic aerosol precursors have been eliminated from the test set. Alternately, the authors may account for these compounds in different

ways. For example, by "not considered functional groups" do the authors mean that other subgroups are used to describe these compounds? I assume that this is the case, since they've included limonene and propylbenzene in Table 2 and a number of aromatics in Table 3. E.g. instead of just "alkenes" or "aromatics", do the methods compared have "=C=", "=CH-", and "=CH2" groups, subdivided further into straight chain or aromatic chains. Or have they excluded two of the classes of compounds which contribute the largest share of secondary organic aerosol formation from their work (aromatics being the largest contributor to anthropogenic SOA, and monoterpenes the largest contributor to biogenic SOA)? If the latter is the case, it lessens the value of the work with respect to the ambient atmosphere, and needs more justification.

Page 18389, equation (15): Given the rigour of the selection process for vapour pressures earlier, the introduction of this equation to allow a larger number of compounds to be compared to "observations" raises some concerns. How accurately are the terms in the equation known, and how does this affect the accuracy of the resulting estimate of sub-cooled liquid vapour pressure? The authors make an attempt earlier to use only direct observations for their experimental data, but here they allow an indirect estimation to be used to generate the data. This needs to be clarified.

Page 18389, line 5: minor issue: could the authors include CAS codes for the species in table 3, hence allowing other researchers to easily investigate the structures in existing databases?

Page 18390, equation (17): not clear if "n" is across different compounds within a class or over all compounds examined in the dataset, from the text.

Page 18390, general comment on section 3.1. If I've understood correctly, the procedure was to find the Tb value giving the closest Pest to Pexp (i.e. minimizing eqn 17), then do stats on the resulting Tb compared to the known Tb? There needs to be a sentence towards the start of this section making this clear.

Page 18391, line 16. The authors treat the relative simplicity of equation 2 as a reason for its use, without giving justification for doing so. It makes sense to select an equation due to it giving smaller standard deviations relative to observations than a more complex equation, but not just because it is a simpler equation. The sentence seems to equate value of the two reasons (smaller errors and simplicity of the equation) as reasons for selecting eqn 2, without justifying the second reason.

Page 18393, lines 5 through 12: While the small sample size invalidates comparisons twixt the figures, there is nothing to prevent the authors from regenerating figure 1 using the subset of the species in figure 2, in order to allow that comparison. Please do so, and discuss.

Page 18395, equation (20): Why use this version of the Antoine equation versus a higher order? The authors mention both extrapolation and interpolation? A higher order would be safer for the latter than for the former. Given the earlier attempts to use only direct experimental data, the use of extrapolation here needs a bit of justification/discussion.

Page 18396, lines 5-7: Did all the methods employed give the same compounds in the Donahue bins, and in the same relative distribution? One might expect not, from the earlier discussion, but if so, it would argue that the condensation might be less sensitive to the method employed than might be expected from the earlier discussion.

Interactive comment on Atmos. Chem. Phys. Discuss., 9, 18375, 2009.