

Interactive comment on “Impact of structure on the estimation of atmospherically relevant physicochemical parameters” by Gabriel Isaacman-VanWertz and Bernard Aumont

Anonymous Referee #1

Received and published: 18 November 2020

This is a rigorously and meticulously conducted study about an interesting and important topic. It is also exceptionally well written and beautifully illustrated. I should be accepted after formulating some very general statements more cautiously (see below for details).

Nevertheless, I can't help feeling somewhat disappointed by the study. The reason is that the study (and by inference a part of the atmospheric science community) continues to rely to a very large extent on property prediction methods that are arguably unsuited to the task of predicting phase partitioning of atmospheric oxidation products. The semi-volatile compounds, for which the secondary organic aerosol (SOA) com-

[Printer-friendly version](#)

[Discussion paper](#)



munity requires vapor pressure and Henry's law constant estimates, invariably have multiple functional groups and as such are unquestionably outside of the applicability domain (AD) of the structure activity relationships (SARs) that this community is using. The more functional groups there are on a molecule, the more opportunities there are for these functional groups to interact. Clearly, these intramolecular interactions influence the ability of molecules to interact with other molecules (such as those in SOA or an aqueous aerosol phase). Yet, most of the SARs used in this study were calibrated using empirical data for compounds with very few functional groups, assume additivity of group contributions, and ignore intramolecular interactions. The authors and the community at large in fact is largely aware of these shortcomings, may indirectly even acknowledge them (e.g. lines 50-53, lines 305-306), but then ignores them and keeps using those SARs (out of inertia?). This may have been justifiable as long as there were no alternatives, but that is a stance that is increasingly more difficult to defend in light of the availability of alternative approaches (Wania et al. Novel methods for predicting gas-particle partitioning during the formation of secondary organic aerosol. *Atmos. Chem. Phys.*, 14, 13189-13204, 2014, Wang et al. Uncertain Henry's law constants compromise equilibrium partitioning calculations of atmospheric oxidation products. *Atmos. Chem. Phys.*, 17, 7529-7540, 2017).

The study does not formally test whether the compounds for which predictions are made are within the AD of the prediction methods; the study neither seeks to quantify the prediction error that is incurred when applying a method to a compound outside of its AD (although both should be possible). Instead it attempts to quantify that error by comparing the results of several prediction methods, that are all leveraged well beyond their AD. This approach is unlikely to give a reasonable estimate of the true prediction uncertainty, because these methods are similar in terms of assuming additivity of group contributions and ignoring intramolecular interactions, i.e. the likelihood that they are all similarly biased is considerable. In what really amounts to circular reasoning, three of the vapor pressure estimation methods are assumed to be "the best available" (line 461) by virtue of the similarity of their predictions (line 463). Predictions that are

[Printer-friendly version](#)[Discussion paper](#)

deemed outliers are simply disregarded in this analysis of uncertainty (e.g. Myrdal and Yalkowsky and EPI, line 454).

I don't suggest that the authors revise their current manuscript by including in the analysis one or the other of the alternative prediction techniques (although they should certainly use those techniques in future endeavors). One recommendation, however, is to make the SMILES codes of the 38,594 atmospherically relevant species representing approximately 1,200 formulae available as a supplementary data file, so that others may be able to conduct a complementary analysis with property prediction techniques, whose ADs actually do comprise those multi-functional compounds.

Line 51, 617: The habit of referring to the Henry's law constant as "solubility" is very unfortunate. The solubility is more appropriately the concentration in a solvent (e.g. water) at saturation. The Henry's law constant (as used in the atmospheric chemistry community) is really the ratio of the solubilities in a solvent and in air (see for example: Cole, J. G., D. Mackay. Correlating environmental partitioning properties of organic compounds: The three-solubility approach. Environ. Toxicol. Chem. 2000, 19, 265-270.) Here on line 50, the reason for the larger divergence in the HLC with increasing number of functional groups occurs because of the lower volatility and not because of a higher water solubility (i.e. the reason for the poorer performance for multi-functional compounds is the same for vapor pressure and HLC).

Incidentally, that is also the reason for the linear relationship between vapor pressure and HLC (Fig. 5a). (However, one has to be very cautious, because the set of compounds investigated here is biased towards chemicals with multiple, polar functional groups, whose activity coefficients in water may not range over a very large range. Large, non-polar chemicals, which are missing from this dataset, would presumably not fit such regressions very well, because of their very high activity coefficients (low water solubilities) in the aqueous phase.)

Line 413: More precisely, this should say: "to understand typical differences in pre-

[Printer-friendly version](#)[Discussion paper](#)

dicted vapor pressure between isomers”. As argued above, the applied SARs cannot credibly claim to capture the real differences in the properties of the isomers of multi-functional compounds.

Line 438: General statements like this need to be qualified: “that predicted vapor pressures of the isomers of atmospheric oxidation products differ by “

Line 498: Statement such as this (“the impact of structure is less than variability in estimation methods”) are really only valid in the context of the very specific assumptions that the authors made. Would they still be true if the Myrdal and Yalkowsky and EPI predictions had not been excluded from the analysis of variability in estimation methods?

Line 604: Another one of these statement that cannot be made with such generality. Just because two estimation methods, when applied to a particular type of compound (outside of their AD!), agree within two orders of magnitude of each other, you cannot conclude that HLC can be estimated to within 2 log units.

Line 704-707: This is an extremely important qualifying statement. It may be advisable to add: - a clarification that these conclusions are also only based on a very particular type of organic compound, namely atmospheric oxidation products with a fairly limited number of functional group types. (This study even noted substantial differences between the atmospheric oxidation products of different precursor VOCs highlighting that the type of compound plays an important role in this regard.) - a sentence that it is entirely possible that these conclusion could be found to not be valid when a different set of prediction methods (or a different type of compounds) were to be employed, in particular if prediction method were used that do not need to be applied well beyond the AD (somewhat euphemistically termed “extrapolating” SARs on line 715).

Line 732-734, 746-747: Again, I commend the authors for including this qualifying statement.

[Printer-friendly version](#)[Discussion paper](#)

Line 97: “less exact than structure-based estimation”

Line 232: “EPI” has not been defined as this stage in the manuscript.

Line 435: In scientific writing the plural of “formula” should be “formulae” and not “formulas”.

Line 430: “isomer pairs”

Line 516: Rephrase: “The above analysis indicates that vapor pressures are sufficiently different between isomers that . . .”

Line 520: The phrase is not comprehensible: “this approach cannot be more precise that than the impacts of structure on a property”?

Line 567: “more accurate than the actual variability”

Line 614: Delete “vapor”

Line 755: “SARs” not “SARS”

Interactive comment on Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2020-1038>, 2020.

Printer-friendly version

Discussion paper

