Response to Comments by Anonymous Referee #1 on "Uncertain Henry's Law Constants Compromise Equilibrium Partitioning Calculations of Atmospheric Oxidation Products" by Chen Wang et al.

This manuscript addresses and explores uncertainty in parameters for modeling phase partitioning of atmospheric organic compounds, a critical outstanding source of uncertainty in current atmospheric chemical models. The authors calculate partitioning coefficients between the vapor phase and both aqueous and organic condensed phases using three different approaches built on different underlying methodologies. Discrepancies between these parameter estimation techniques are discussed and used to identify current critical gaps in understanding. In addition, the results of each approach are explored in the context of ambient conditions, and the authors demonstrate that differences between approaches significantly change the expected phase of many organic compounds in the atmosphere. This work is generally a valuable step toward understanding and eventually addressing current shortcomings in atmospheric partitioning models, and this reviewer recommends publication with only minor revisions.
Response: Thanks for the comments.

General comments:
1) By providing the calculated results for all species, this work significantly advances future modeling. The Excel spreadsheet seems to be corrupt, though, or at least did not work on my computer. Please fix this, and perhaps provide the data in a more portable form as well, such as CSV.
Response: Thanks for the reviewer's suggestions. We have updated the excel spreadsheet and added a csv file in the supporting information.

2) Additional detail about the three estimation approaches, in particular SPARC and ppLFER, should be provided in the methods. Throughout the manuscript the reader has to sort of piece together everything that goes into these two approaches. They should instead be given more explanation in the Methods section.
Response: We have rewritten lines 122-127 in the Method section to include more details about the ppLFER method: "In brief, ppLFERs are developed by performing a multi-linear regression of experimental *K* values against compound specific solute descriptors (Endo and Goss, 2014). These descriptors represent a solute's hydrogen-bond acidity (*A*), hydrogen-bond basicity (*B*), dipolarity/polarizability (*S*), McGowan volume (cm^3/mol) divided by 100 (*V*), excess molar refraction (*E*), and logarithmic hexadecane-air partitioning constant at 25°C (*L*). In this study, solute descriptors for the 3414 compounds were predicted with ABSOLV (ACD/Labs, Advanced Chemistry Development, Inc., Toronto, Canada). The regression coefficients in ppLFERs are denoted by *a*, *b*, *s*, *v*, *e*, and *l*; *c* is the regression constant. The ppLFER for air-water partitioning was taken from (Goss, 2006):

$$\log K_{W/G} = c + aA + bB + sS + vV + eE$$

whereas ppLFERs for four different organic aerosol were taken from (Arp et al., 2008);

$$\log K_{Aerosol/G} = c + aA + bB + sS + vV + lL$$

As described in Wania et al. (2014), the average of the four $K_{Aerosol/G}$ was compared with the $K_{WIOM/G}$ predicted by the other two methods." We have also added the reference Endo and Goss (2014) which gives a comprehensive introduction to ppLFERs: "Endo, S., Goss, K.-U., Applications of polyparameter linear free energy relationships in environmental chemistry, Environmental Science and Technology, 48, 12477-12491, 2014. "

We have further added more details on the SPARC method on line 127: "SPARC is a commercial web-based calculator for prediction of physical chemical properties from molecular structure developed by the US Environmental Protection Agency (Hilal et al., 2004). The predictions of $K_{W/G}$ and $K_{WIOM/G}$ are based on solvation models in SPARC that describe the intermolecular interaction between different molecules (solute and solvent), including dispersion, induction, dipole-dipole, and H-bonding interactions, which are developed and calibrated with experimental data (Hilal et al., 2008)."

3) The section on "Comparison between Different Prediction Methods" focuses on MD and MAD, but this somewhat masks the true scope of uncertainty. For instance, in Table 1, these metrics suggest the $K_{W/G}$ comparison between ppLFER and SPARC is not much different than the $K_{WIOM/G}$ comparison except for >5 functional groups. Claims by the author to the contrary are somewhat overstated. From the breakdown by functional groups and from Figure 1, though, it is clear there are some extreme or at least more varied cases. It seems relevant not only to ask "what is the average difference?" but also to ask "what is the probability that these two methods differ substantially?" Including as an additional figure a distribution (or cumulative distribution) of differences would help answer this question by showing not only average difference (the center of the distribution), but also the range of differences (the width and range of the distribution), and would strengthen the author's claims that there is a substantial difference in the uncertainty of these parameters.
Response: We have added a figure to the supporting information (see Figure S1 at the end of this document) with plots showing the frequency of the discrepancies for predictions between any two prediction methods and added the following sentence on line 167: "Figure S1 in the supporting information illustrates the frequency of the discrepancies between different pairs of predicted log $K_{WIMO/G}$ and log $K_{W/G}$ values." The figure numbers in the manuscript and Supporting information have been changed accordingly.

4) In discussing atmospheric implications of different prediction methods, an important metric is the number (or fraction) of compounds that are in a different phase with different prediction methods, not just the number in each phase with each method as in Table 2. For instance, how many compounds that are condensed with ppLFER that "volatilize" with COSMOtherm? This would highlight the implications and importance of the differences.
Response: This can be evaluated by comparing the fraction of a certain compound in the gas phase, i.e. whether or not it is present mostly in the condensed phase, predicted by different method under certain conditions (WIOM phase and liquid water content). This is illustrated in the partitioning space plots in Figure 3.

In addition, we calculated how many (percentage) of the compounds change their preferred phase when a different estimation method is used. The threshold used was 50 % in the gas phase, i.e. if a compound is less than 50 % in the gas phase it is predominantly in the condensed phase. The number of compounds changing from being predominantly present in the gas phase to being predominantly in the condensed phase when a different method is used is summarized in Table S1. The following sentence has been added to the manuscript on line 448: "Table S1 in the supporting information summarizes the number and percentage of compounds that change their partitioning between gas and condensed phase under different atmospheric conditions when a different prediction method is used. Depending on the scenarios, a total of 2.0 % up to 34 % of the 3414 compounds have a different dominant phase when using a different prediction method. This change is larger for the cloud scenarios and much lower for the aerosol scenarios especially if the aerosol contains no water."

Table S1 Number (percentage) of compounds that change from predominant partitioning to gas phase to predominant partitioning to the condensed phase(s) under different atmospheric conditions when a different prediction method is used

(a) aerosol (LWC=10 μg/m^3, OM=10 μg/m^3)			**(c) aerosl without water (LWC=0 μg/m^3, OM=10 μg/m^3)**		
from ppLFER to SPARC	from SPARC to ppLFER	total change	from ppLFER to SPARC	from SPARC to ppLFER	total change
107 (3.1%)	195 (5.7%)	302 (8.8%)	146 (4.3 %)	11 (0.3%)	157 (4.6%)
from ppLFER to COSMOtherm	from COSMOtherm to ppLFER	total change	from ppLFER to COSMOtherm	from COSMOtherm to ppLFER	total change
17 (0.5%)	481 (14.1 %)	498 (14.6%)	26 (0.8%)	41 (1.2%)	67 (2.0%)
from SPARC to COSMOtherm	from COSMOtherm to SPARC	total change	from SPARC to COSMOtherm	from COSMOtherm to SPARC	total change
12 (0.4%)	388 (11.4%)	400 (11.7%)	14 (0.4%)	164 (4.8%)	178 (5.2%)
(b) cloud (LWC=0.3 g/m^3, OM=10 μg/m^3)			**(d) cloud non-phase separated (LWC=0.3 g/m^3, OM=0 μg/m^3)**		
from ppLFER to SPARC	from SPARC to ppLFER	total change	from ppLFER to SPARC	from SPARC to ppLFER	total change
166 (4.9%)	207 (6.1%)	373 (10.9%)	167 (4.9%)	206 (6.0%)	373 (10.9%)
from ppLFER to COSMOtherm	from COSMOtherm to ppLFER	total change	from ppLFER to COSMOtherm	from COSMOtherm to ppLFER	total change
46 (1.3%)	1103 (32.3 %)	1149 (33.7%)	47 (1.4%)	1111 (32.5%)	1158 (33.9%)
from SPARC to COSMOtherm	from COSMOtherm to SPARC	total change	from SPARC to COSMOtherm	from COSMOtherm to SPARC	total change
40 (1.2%)	1056 (30.9 %)	1096 (32.1%)	40 (1.2%)	1065 (31.2%)	1105 (32.4%)

*The column "total change" indicates total number (percentage) of compounds that have different predicted dominant phases using different methods.

Specific comments:

p. 3 line 49: "number of organic species in the atmosphere is in the hundreds of thousands." Please cite your source, as these numbers often vary in the literature between thousands, tens of thousands, and hundreds of thousands, but probably only if including constitutional isomers with the same functional groups in different positions.

Response: This sentence (line 49) has been changed: "Furthermore, there are many thousand organic species in the atmosphere (Hallquist et al., 2009); the number is even higher when considering their isomers." The following reference has been added: Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prevot, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, Atmospheric Chemistry and Physics, 9, 5155-5236, 2009.

p. 3 line 58: Suggest including a reference to Compernolle et al., doi:10.5194/acp-11-9431-2011, which also explores this issue in some detail, specifically comparing various v.p. estimation methods used in this field

Response: The reference suggested by the reviewer has been included: "Compernolle, S., Ceulemans, K., and Muller, J. F.: EVAPORATION: a new vapour pressure estimation method for organic molecules including non-additivity and intramolecular interactions, Atmospheric Chemistry and Physics, 11, 9431-9450, 10.5194/acp-11-9431-2011, 2011."

p. 4 line 76-78: Here and throughout the paragraph, it may be worth noting the expected uncertainties in some or all of these methods. The Hodzic approach suffers from fairly large scatter in the c*-H_{eff} trend. The authors also mention the cross comparison of GROMHE SPARC and HENRYWIN, and later cite a similar such comparison by Isaacman-VanWertz et al., but don't mention the results of these comparisons here (several orders of magnitude discrepancy). This paragraph would better motivate the work by giving a quantitative discussion of previous estimates of variation across methods.

Response: We have added the following discussion on line 85 on the quantitative performance of the estimation methods for the Henry's constant: "Even for the relatively simple molecules for which experimental evaluation data exist, these methods have considerable uncertainties. Raventos-Duran et al. (2010) reported Root Mean Square Errors (RMSE) of 0.38, 0.61, and 0.73 log units for Henry's constants predicted by GROMHE, SPARC and HENRYWIN, respectively. The ppLFER developed by Goss (2006) has a RMSE of 0.15 log units for the 217 compounds used for calibration. The error can be expected to be much larger for molecules that either are not part of the calibration (GROHME, ppLFER) or are more complex. For a compound with multiple functional groups, Isaacman-VanWertz et al. (2016) found discrepancies in predicted Henry's law constant of several orders of magnitude. Hodzic et al. (2014)'s method of estimating the Henry's law constant for atmospheric oxidation products of different precursors also has uncertainties of several orders of magnitude."

p. 5 line 97: Why not use all non-radical species in the MCM? Or is 3414 all of them? If not, what was excluded and why?

Response: The 3414 compounds include all of them.

p. 5 line 107: Should be "units" instead of "unit"

Response: "unit" on line 107 and 108 are changed to "units".

p. 6 line 122: Based on comments throughout the manuscript, it sounds like ppLFER includes some empirical calibrations- please elaborate a bit more on this approach.

Response: Details of this method have been added to the Method section. Please refer to response to an earlier comment.

p. 6 line 123: What is a "solute descriptor"? Please define

Response: Details of this method have been added to the Method section. Please refer to response to an earlier comment.

p. 6 lines 130-131: See general comment 2. A lot more information is provided about COSMOtherm than ppLFER or SPARC. Please provide a one-sentence description of what approach to these calculations SPARC takes

Response: A detailed description of the SPARC method has been added in the Method section.

p. 7 lines 155, 162: It is a little confusing to including the Hodzic ranges in both their units and K_CP/G units. Consider sticking to the latter.

Response: We have changed lines 155-159 as follows: "Hodzic et al. (2014) predicted a log $K_{WIOM/G}$ in the range of approximately 0 and 20 at 25 °C (see conversion between C* and $K_{WIOM/G}$ in the supporting information) for oxidation products of different VOCs (including n-alkanes, benzene, toluene, xylene, isoprene and terpenes), i.e. their data set included higher $K_{WIOM/G}$ values than those generated here, even though $K_{WIOM/G}$ values are lower at higher temperature."

Lines 162-164 have been changed to: "Hodzic et al. (2014) predicted a log $K_{W/G}$ in the range of -2.6 and 17.4 at 25°C (see conversion to K_H in the supporting information)."

p. 7 line 158: To add clarity, consider reminding the reader of physical meaning when using statements like "higher $K_{WIOM/G}$", such as adding a parenthetical "(lower volatility)".
Response: We added "(indicating generally lower volatility)" after "higher $K_{WIOM/G}$".

p. 7 line 167: It would be worth pointing out early in this section that agreement between methods does not confirm or disconfirm accuracy. An easy first conclusion from Figure 1 is that COSMOtherm is just way off in K_W/G since the others agree. This is a conclusion the authors thoroughly discuss and debunk later, but it may help to guide readers away from this conclusion in the first place
Response: The following sentence has been added on line 167: "This discrepancy only indicates the agreement between any two predictions with little indication of the accuracy of the prediction for reasons discussed later."

p. 10 line 208: Again, consider adding Compernolle et al. to this citation.
Response: This reference has been added on line 206.

p. 10 line 224: It overstates the data to claim that "K_WIOM/G is almost always smaller than one log unit". Of the 21 functional group comparison "bins" in table 1, 5 have MAD above 1 log unit, and another 4 have MAD between 0.9 and 1. So 20-40% of the bins fall outside or nearly outside this claim.
Response: The discrepancy in predicted log $K_{WIOM/G}$ is smaller than 1 log unit for 64 % (ppLFER vs. SPARC), 66 % (COSMOtherm vs. SPARC) and 75% (COSMOtherm vs. ppLFER) of the 3414 compounds. We changed "almost always" on line 224 to "mostly (and on average)".

p. 13 line 282: "Partition" should be "partitioning"
Response: Changed accordingly.

p. 13 line 293-295: Here and below, the authors suggest that a lot of the issue with ppLFER lies in the limitations of solute description from ABSOLV, but do not discuss a means for improving this descriptors. What data would the authors need for this? This should be discussed, because if there is no way to get improved data, then this is an inherent limitation of ppLFER, or on the other hand it may be trivial to improve ppLFER in future work.
Response: A detailed description on how to empirically determine solute descriptors for organic substances is given in Endo and Goss (2014). We have added the following sentence: "While the use of measured solute descriptors therefore would likely greatly improve the ppLFER prediction (Endo and Goss, 2014), those are unlikely to become available for atmospheric oxidation products."

p. 14 line 326: Again, a little confusing to switch between K_{CP}, v.p. and C* in discussion

Response: We added “and C*, i.e. underestimating the volatility of the organic compounds” after “vapor pressures (P_L)” on line 320 to help the readers to understand the discussion.

p. 15 line 330: Define or remove P_L
Response: P_L is vapor pressure and has been defined when it appeared first on line 320.

p. 15 line 351-352: Transition to the bulleted list is awkward. Change to: "However, we can infer that: - the fact.... - the generally...."
Response: Changed accordingly.

p.15 line 353-354: Again, it overstates the data to claim "$K_{WIOM/G}$ that are on average within one order of magnitude for all studied compounds" particular when including the claim including highly oxygenated multifunctional organic compounds," which differ by 1.5-2 orders of magnitude between COSMOtherm and the others
Response: “including highly oxygenated multifunctional organic compounds" on line 354 has been changed to “and less than two orders of magnitude for highly oxygenated multifunctional organic compounds”

p. 16 line 359: This is the first mention the ppLFER use real aerosols as a calibration reference. This highlights that information about what exactly goes into ppLFER is spread throughout the manuscript, it should be discussed in much more detail in the methods.
Response: A detailed description has been added in Method section.

p. 19 line 439: See general comment 4. Quantifying the compounds that switch from condensed- to gas-phase between methods would provide more insight into the potential impact on SOA mass. Note that this is different than just the number of compounds in each phase with each method as in Table 2. A compound in the WIOM phase in all 3 methods doesn’t "care" what method is used. Instead, the relevant metric for discussing SOA implications here and throughout the paragraph is changes in phase, in particular changes from condensed- to gas-phase.
Response: Please refer to the response to general comment 4.

Figures 1 and 2: Considering that much of the discussion is comparing difference in $K_{WIOM/G}$ vs. $K_{W/G}$, it would be helpful to keep the top and bottom panels on the same y-axis scale. Also, in the headings of "Y vs X", generally X is on the x-axis and Y is on the y-axis, instead of the opposite used here
Response: Figure 1 has been changed according to the reviewer’s suggestion (see below). Figures S2, S4, S6, S7, S8 in supporting information have also been changed accordingly.

We did not change the scales in Figure 2 for a better illustration of the data since the range of the discrepancy for $K_{W/G}$ is much larger than that for $K_{WIOM/G}$. We added a note under the caption of Figure 2 on line 237 to clarify the differences in the scales: “Note the different scales for different panels.”

Figure 2: Gridlines on the major y-axis ticks would be helpful
Response: We have added gridlines for y-axis ticks in Figure 2.

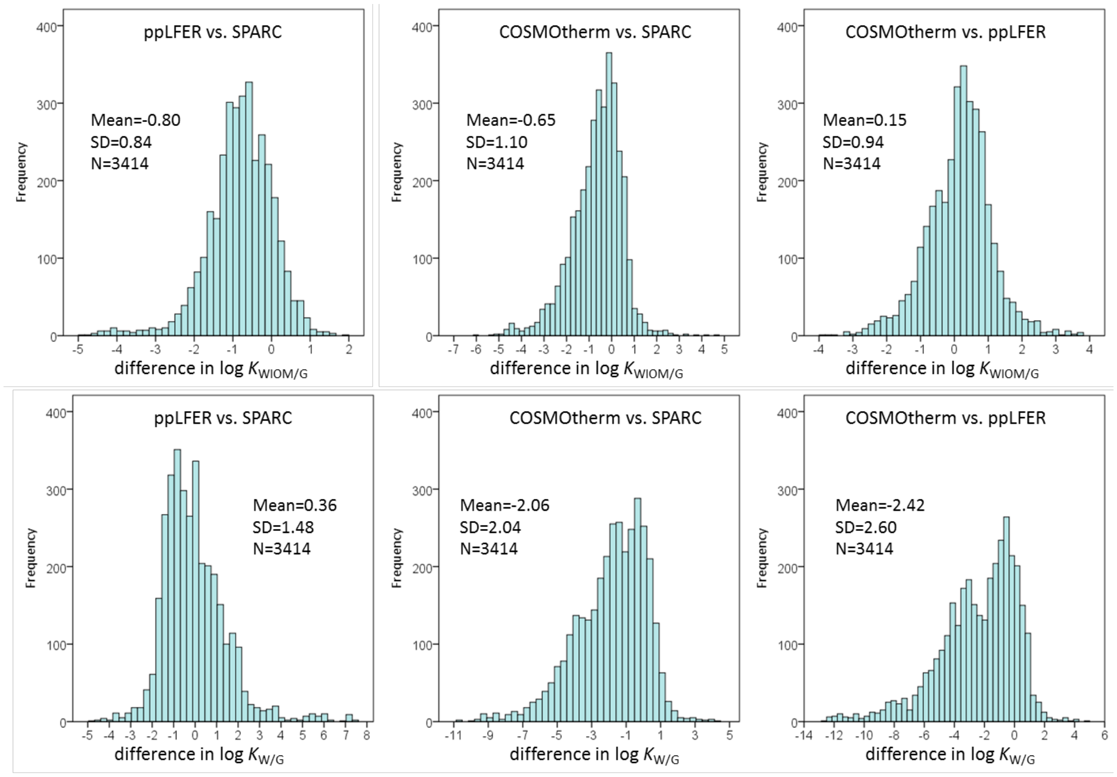
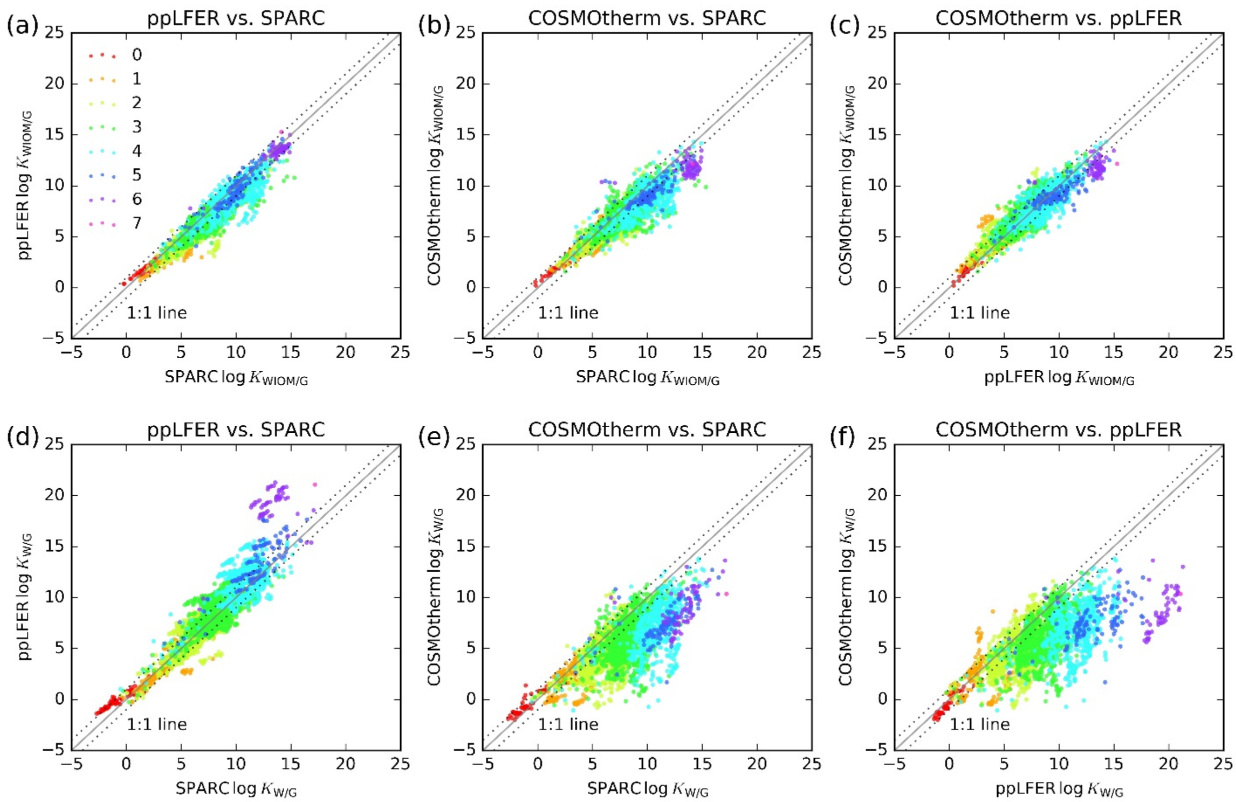
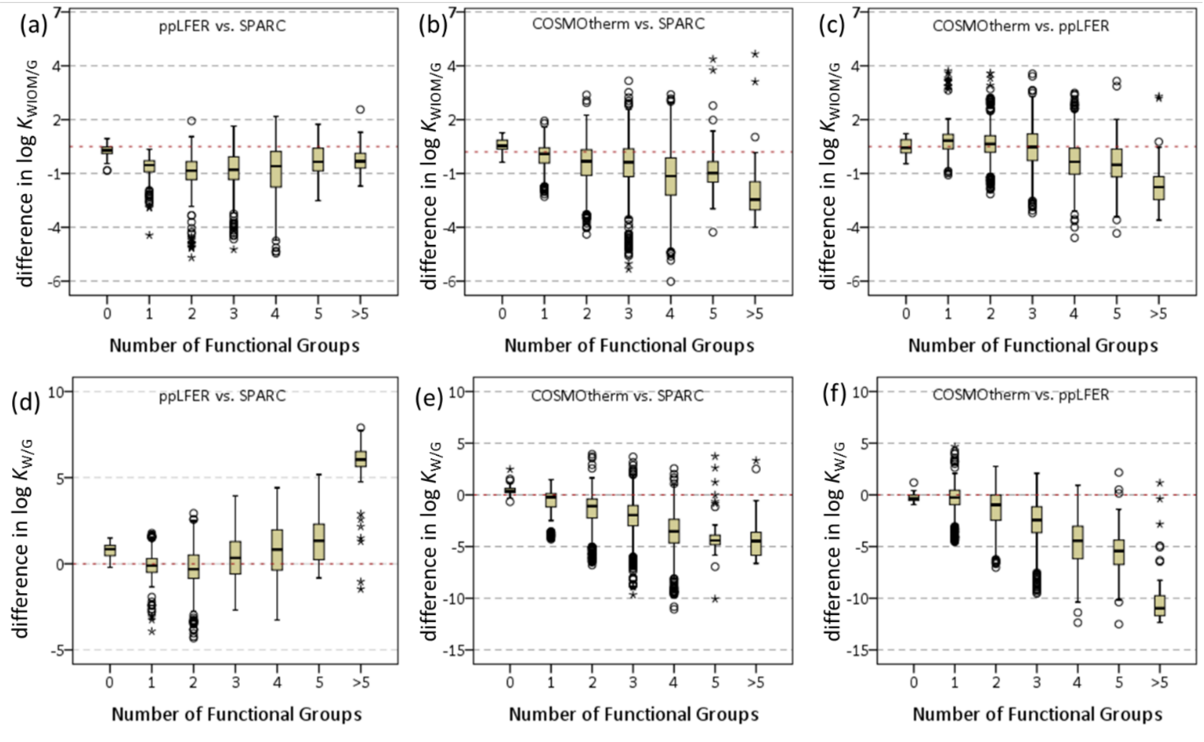


Figure S1 Frequency of discrepancies between different pairs of predictions of log $K_{WIOM/G}$ (top) and log $K_{W/G}$ (bottom).



Revised Figure 1



Revised Figure 2

Response to Comments by Anonymous Referee #2 on "Uncertain Henry's Law Constants Compromise Equilibrium Partitioning Calculations of Atmospheric Oxidation Products" by Chen Wang et al.

This manuscript describes in detail a modeling experiment to determine the best approach to describe partitioning of organic gases (thousands of compounds tested) into the particle-phase's aqueous and organic medium. The authors employ 3 modeling approaches to describe partitioning with a focus on highly oxidized material. The authors also offer comparison and a critique of an approach currently implemented in an atmospheric model based on volatility. The authors make a compelling argument for their main thesis: "The large uncertainty in Kw/g predictions for highly functionalized organic compounds needs to be resolved to improve the quantitative treatment of SOA formation." Predicted organic aerosol amounts in atmospheric models will be highly dependent on and sensitive to the chosen partitioning parameterizations, which are highly uncertain. The authors identify a key knowledge gap.

I recommend the paper for publication provided adequate response and revision to the comments provided below.
Response: Thanks for the comments.

My biggest challenge understanding this paper was Figure 3, which I believe is the most important. Perhaps there is a way to draw in 3 dimensions to make more clear?
It is confusing to have the vertical purple line "without aqueous phase" drawn in the aqueous phase. It is also confusing to just have this scenario for only the ppLFER experiments. Casual readers will not understand what the circled dots in the Figure 3c are.
Response: We have simplified Figure 3 and added two more figures (S13 and S14) in the supporting information to make the figures more understandable (see figures below). The text in the manuscript has been modified accordingly.

Lines 383-387 have been changed to: "The blue dotted lines represent a cloud scenario where LWC is 0.3 g/m^3 and OM is 10 μg/m^3. Figures S13 and S14 in the supporting information show an aerosol scenario without an aqueous phase and a cloud scenario without a separated organic phase because all of the OM is dissolved in the aqueous phase (see also Figure S12 (c) and (d))."

The following sentence was added at the end of line 433 "Those compounds are not sufficiently soluble in water to partition to the cloud and are not sufficiently volatile to be in the gas phase."

"Figure S12" on line 408 and 411 was replaced with "Figure S15".

Line 448 has been changed to: "The number of compounds on the right side of the blue dotted boundary in Figure S13 does not vary substantially with different predictions."

Why do there appear to be 'straight' lines in the dots for all models, most pronounced for 0 and 1 functional groups?
Response: There are no straight lines in the dots in Figure 3 so nothing has been changed.

Page 4, Line 71/72: May an additional reason for the study and importance of VOC oxidation products be that in addition to their higher affinity, they have a great atmospheric abundance?
Response: On line 72, we add "and a great atmospheric abundance."

Figure 2: can the method for ‘possible outlier’ and ‘extreme value’ be explicitly stated here

Response: The “possible outliers”, i.e., the circles, are values that are either 1.5×IQR or more above the third quartile or 1.5×IQR or more below the first quartile, where IQR is the range between the first and third quartile of the boxplot, called interquartile range (IQR). The asterisks or stars are “extreme outliers”, which are either 3×IQR or more above the third quartile or 3×IQR or more below the first quartile.

Editorial: p. 7, Line 159: “value” should be “values”

Response: Changed.

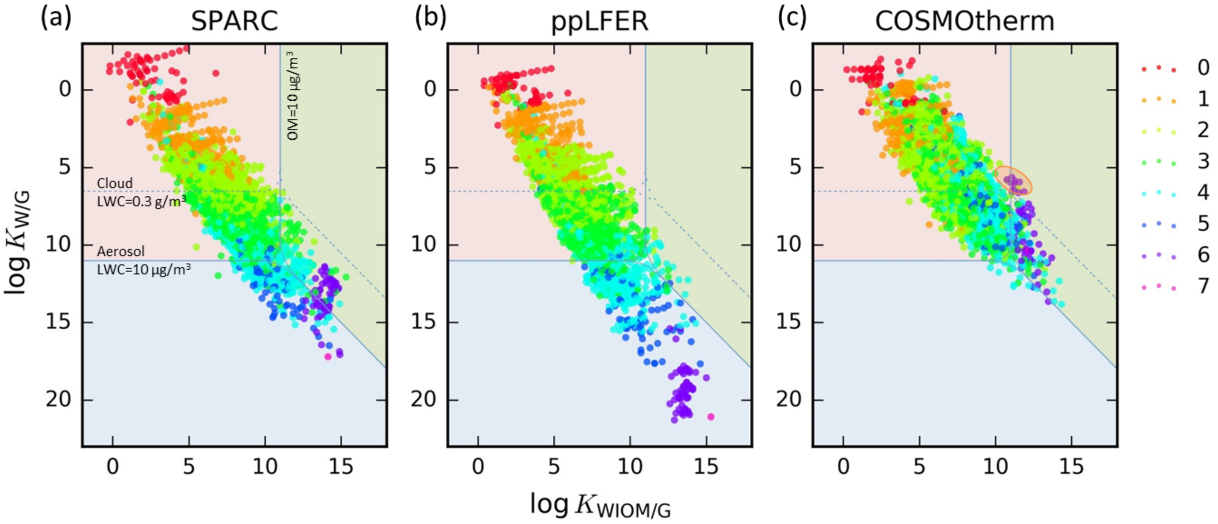


Figure 3 Partitioning space plot, showing in pink, blue and green the combinations of partitioning properties that lead to dominant equilibrium partitioning to the gas, aqueous, and WIOM phases, respectively. The blue solid and dotted lines are boundaries for an aerosol scenario (LWC 10 μg/m^3, 10 μg/m^3 OM) and a cloud scenario (LWC 0.3 g/m^3, 10 μg/m^3 OM), respectively. The differently colored dots indicate the number of functional groups in the molecules.

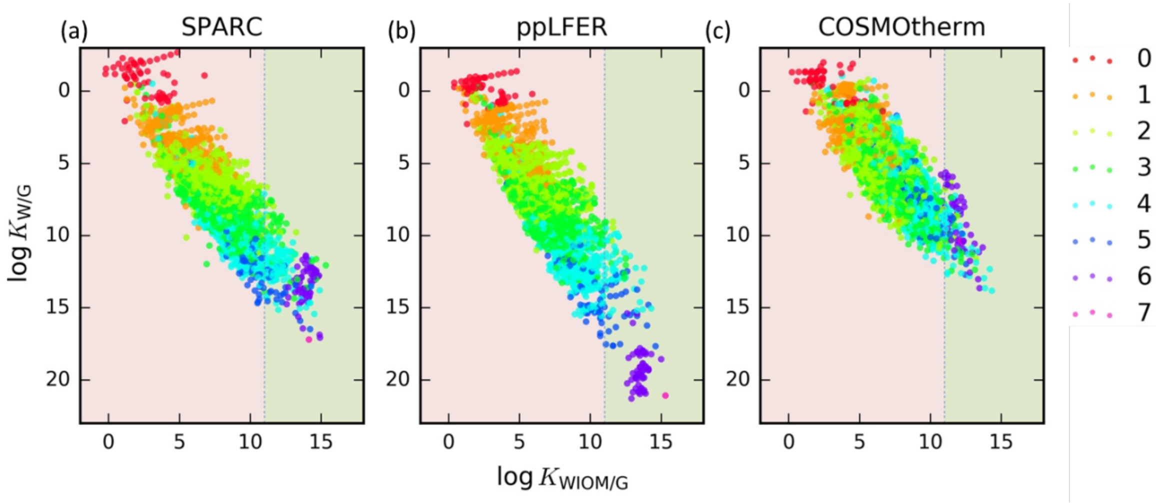


Figure S13 Partitioning space plot, showing in pink and green the combinations of partitioning properties that lead to dominant equilibrium partitioning to the gas and WIOM phases, respectively. The blue dotted lines are boundaries for an aerosol scenario without an aqueous phase (LWC 0 μg/m^3, 10 μg/m^3 OM). The differently colored dots indicate the number of functional groups in the molecules

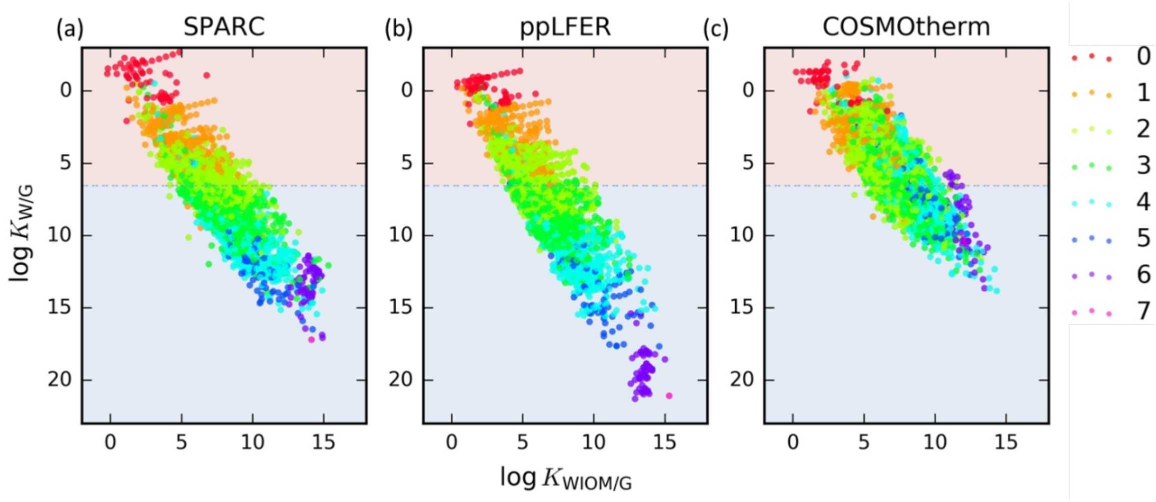


Figure S14 Partitioning space plot, showing in pink and blue the combinations of partitioning properties that lead to dominant equilibrium partitioning to the gas and aqueous phases, respectively. The horizontal blue dashed lines a cloud scenario where LWC is 0.3 g/m^3 and OM 0 μg/m^3. The differently colored dots indicate the number of functional groups in the molecules.

Uncertain Henry's Law Constants Compromise Equilibrium Partitioning Calculations of Atmospheric Oxidation Products

Chen Wang[1], Tiange Yuan[1], Stephen A. Wood[1], Kai-Uwe Goss[2,3], Jingyi Li,[4,5] Qi Ying[4], Frank Wania[1*]

[1]Department of Physical and Environmental Sciences, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON, M1C 1A4, Canada

[2]Department of Analytical Environmental Chemistry, Centre for Environmental Research UFZ Leipzig-Halle, Permoserstraße 15, Leipzig, D-04318, Germany

[3]Institute of Chemistry, University of Halle-Wittenberg, Kurt-Mothes-Straße 2, Halle, D-06120, Germany

[4]Environmental and Water Resources Division, Zachry Department of Civil Engineering, Texas A&M University College Station, TX 77843-3136

[5]Current Address: School of Environmental Science and Engineering, Nanjing University of Information Science & Technology, 219 Ningliu Road, Nanjing 210044, China

[*]Corresponding author: frank.wania@utoronto.ca, +1-416-287-7225

Abstract

Gas-particle partitioning governs the distribution, removal and transport of organic compounds in the atmosphere and the formation of secondary organic aerosol. The large variety of atmospheric species and their wide range of properties make predicting this partitioning equilibrium challenging. Here we expand on earlier work and predict gas-organic and gas-aqueous phase partitioning coefficients for 3414 atmospherically relevant molecules using COSMOtherm, SPARC and poly-parameter linear free energy relationships. The Master Chemical Mechanism generated the structures by oxidizing primary emitted volatile organic compounds. Predictions for gas-organic phase partitioning coefficients ($K_{\mathrm{WIOM/G}}$) by different methods are on average within one order of magnitude of each other, irrespective of the numbers of functional groups, except for predictions by COSMOtherm and SPARC for compounds with more than three functional groups, which have a slightly higher discrepancy. Discrepancies between predictions of gas-aqueous partitioning ($K_{\mathrm{W/G}}$) are much larger and increase with the number of functional groups in the molecule. In particular, COSMOtherm often predicts much lower $K_{\mathrm{W/G}}$ for highly functionalized compounds than the other methods. While the quantum-chemistry based COSMOtherm accounts for the influence of intramolecular interactions on conformation, highly functionalized molecules likely fall outside of the applicability domain of the other techniques, which at least in part rely on empirical data for calibration. Further analysis suggests that atmospheric phase distribution calculations are sensitive to the partitioning coefficient estimation method, in particular to the estimated value of $K_{\mathrm{W/G}}$. The large uncertainty in $K_{\mathrm{W/G}}$ predictions for highly functionalized organic compounds needs to be resolved to improve the quantitative treatment of SOA formation.

Introduction

Volatile organic compounds (VOCs) emitted to the atmosphere are oxidized to form secondary products. These products tend to be more oxygenated, less volatile and more water-soluble than their parent compounds, and thus have higher affinity for aerosol particles and aqueous droplets. Equilibrium partitioning coefficients are often needed to assess the distribution of these oxidized compounds among different phases in the atmosphere such as aerosol particles, fog and cloud droplets. In particular, the partitioning between gas and organic phase and between gas and aqueous phase is required for the evaluation of an organic compound's contribution to secondary organic aerosol (SOA) formation, its transport, removal and lifetime. Experimentally determined partitioning coefficients are rarely available for the oxidation products of VOCs due to the difficulties in making the measurements and obtaining chemical standards. Furthermore, there are many thousand organic species in the atmosphere (Hallquist et al., 2009); the number is even higher when considering their isomers. Their gas-particle partitioning is therefore usually predicted. Reliable estimation methods for gas-organic and gas-aqueous partitioning should be applicable to a wide range of organic compounds, especially to multifunctional species generated during the multi-step atmospheric oxidation of precursor VOCs.

Current approaches for predicting partitioning into non-aqueous organic aerosol phases almost exclusively rely on predictions of vapor pressure. These predictions have large uncertainties; comparison among different vapor pressure prediction methods suggest increasing discrepancies with increasing numbers of functional groups in an organic compound (Valorso et al., 2011;Barley and McFiggans, 2010;McFiggans et al., 2010;Compernolle et al., 2011). This uncertainty matters, because it is the multi-functional oxidation products that can occur in either gas or condensed phases in the atmosphere. Instead of relying on predictions for vapor pressures, Wania et al. (2014) proposed using three alternative methods for direct gas-particle partitioning prediction: poly-parameter linear free energy relationships (ppLFERs), the on-line calculator of SPARC Performs Automated Reasoning in Chemistry (SPARC) and the quantum-chemistry based program COSMOtherm. Wania et al. (2014) found that partitioning coefficients predicted for the oxidation products of n-alkanes are within one order of magnitude, and mutual agreement does not deteriorate with increasing number of functional groups. Because of the relatively small number of oxidation products in that study, the reliability of these prediction methods for other organic compounds requires further evaluation.

While more experimental data exist for the Henry's law constant of atmospherically relevant compounds than gas-organic phase partitioning coefficients (Sander, 2015), data are not usually available for VOC oxidation products, which potentially have a higher affinity for atmospheric aqueous phases and a great atmospheric abundance. Currently available prediction methods for the air-water partitioning coefficient include GROup contribution Method for Henry's law Estimate (GROMHE) (Raventos-Duran et al., 2010), SPARC (Hilal et al., 2008), HENRYWIN in EPI suite (US EPA, 2012), and ppLFERs (Goss, 2006). Sander (2015) provides a more comprehensive list of websites as well as quantitative structure-property relationships for Henry's law constants. COSMOtherm can also predict gas-aqueous phase partitioning of organic compounds, including VOC oxidation products (Wania et al., 2015). Though many different methods are available for Henry's law constant prediction, they have not been systematically evaluated for a large set of organic compounds of atmospheric relevance. An exception is the comparison of GROMHE, SPARC and HENRYWIN predictions for 488 organic compounds bearing functional groups of atmospheric relevance (Raventos-Duran et al., 2010).

Even for the relatively simple molecules for which experimental evaluation data exist, these methods have considerable uncertainties. Raventos-Duran et al. (2010) reported Root Mean Square Errors (RMSE) of 0.38, 0.61, and 0.73 log units for Henry's constants predicted by GROMHE, SPARC and HENRYWIN, respectively. The ppLFER developed by Goss (2006) has a RMSE of 0.15 log units for the 217 compounds used for calibration. The error can be expected to be much larger for molecules that either are not part of the calibration (GROHME, ppLFER) or are more complex. For a compound with multiple functional groups, Isaacman-VanWertz et al. (2016) found discrepancies in predicted Henry's law constant of several orders of magnitude. Hodzic et al. (2014)'s method of estimating the Henry's law constant for atmospheric oxidation products of different precursors also has uncertainties of several orders of magnitude.

The objective of this paper was to compare and evaluate gas-particle partitioning predictions for a large number of organic compounds of atmospheric interest using ppLFER (in combination with ABSOLV-predicted solute descriptors), SPARC and COSMOtherm. While all three methods are able to estimate both gas-organic and gas-aqueous partitioning, they are based on different principles: ppLFERs are empirically calibrated multiple linear regressions, SPARC contains solvation models based on fundamental chemical structure theory (Hilal et al., 2004), and COSMOtherm combines quantum chemistry with statistical thermodynamics (Klamt and Eckert, 2000). This study thus expands earlier work (Wania et al., 2014) to a much larger number of compounds and to aqueous phase partitioning. As such, it includes quantum-chemistry based predictions for an unprecedented number of atmospherically relevant compounds.

Method

The Master Chemical Mechanism (MCM v3.2, *http://mcm.leeds.ac.uk/MCM*) a near-explicit chemical mechanism was used to generate 3414 non-radical species through the multi-step gas phase oxidation of 143 parent VOCs (methane + 142 non-methane VOCs). Reactions of the parent VOCs with O_3, OH and NO_3 are included in the MCM mechanism whenever such reactions are possible. The details about the studied compounds are given in the supporting information (Excel spreadsheet), including the compounds' MCM ID, SMILES, precursors (i.e. the parent VOC), molecular weight, molecular formula, elements, generation of oxidation, number and species of functional groups, O:C ratio, and average carbon oxidation state ($\overline{OS}_{\mathrm{C}}$) (Kroll et al., 2011).

Three prediction methods are used to estimate the equilibrium partitioning coefficients between a water-insoluble organic matter phase (WIOM) and the gas phase ($K_{\mathrm{WIOM/G}}$) at 15 °C in units of m^3 (air)/m^3 (WIOM) as well as the equilibrium partitioning coefficients between water and gas phase ($K_{\mathrm{W/G}}$) at 15 °C in units of m^3 (air)/m^3 (water). The two partitioning coefficients are defined as:

$$K_{\mathrm{WIOM/G}} = C_{\mathrm{WIOM}}/C_{\mathrm{G}} \qquad (1)$$

$$K_{\mathrm{W/G}} = C_{\mathrm{W}}/C_{\mathrm{G}} \qquad (2)$$

C_{WIOM}, C_W and C_G (mol/m^3) are equilibrium concentrations of an organic compound in WIOM, water, and gas phase, respectively. Partitioning between gas and aqueous phase can be significantly influenced by the presence of inorganic salts (i.e. the salt effect) (Endo et al., 2012;Wang et al., 2016;Wang et al., 2014;Waxman et al., 2015), the hydration of carbonyls (Ip et al., 2009) and the dissociation of organic acids (Mouchel-Vallon et al., 2013), particularly in the aqueous phase of aerosols. However, in this study only the partitioning between gas and pure water, i.e. the Henry's law constant, is predicted, and no hydration, salt effect or acid dissociation is considered. Conversion of partitioning coefficients $K_{W/G}$ to Henry's constant (K_H) in units M/atm or $K_{WIOM/G}$ to saturation concentration (C^*, µg/m^3) is provided in the supporting information.

Wania et al. (2014) describe each prediction method in detail. In brief, ppLFERs are developed by performing a multi-linear regression of experimental K values against compound specific solute descriptors (Endo and Goss, 2014). These descriptors represent a solute's hydrogen-bond acidity (A), hydrogen-bond basicity (B), dipolarity/polarizability (S), McGowan volume (cm^3/mol) divided by 100 (V), excess molar refraction (E), and logarithmic hexadecane-air partitioning constant at 25°C (L). In this study, solute descriptors for the 3414 compounds were predicted with ABSOLV (ACD/Labs, Advanced Chemistry Development, Inc., Toronto, Canada). The regression coefficients in ppLFERs are denoted by a, b, s, v, e, and l; c is the regression constant. The ppLFER for air-water partitioning was taken from (Goss, 2006):

$$\log K_{W/G} = c + aA + bB + sS + vV + eE \qquad (3)$$

whereas ppLFERs for four different organic aerosol were taken from (Arp et al., 2008);

$$\log K_{Aerosol/G} = c + aA + bB + sS + vV + lL \qquad (4)$$

As described in Wania et al. (2014), the average of the four $K_{Aerosol/G}$ was compared with the $K_{WIOM/G}$ predicted by the other two methods. SPARC is a commercial web-based calculator for prediction of physical chemical properties from molecular structure developed by the US Environmental Protection Agency (Hilal et al., 2004). The predictions of $K_{W/G}$ and $K_{WIOM/G}$ are based on solvation models in SPARC that describe the intermolecular interaction between different molecules (solute and solvent), including dispersion, induction, dipole-dipole, and H-bonding interactions, which are developed and calibrated with experimental data (Hilal et al., 2008). For the calculations of $K_{\mathrm{WIOM/G}}$ by SPARC and COSMOtherm, the phase WIOM is represented by the surrogate structure "B" as proposed by Kalberer et al. (2004) and adopted previously by Arp and Goss (2009) and Wania et al. (2014). SPARC calculations were carried out using the on-line calculator (http://archemcalc.com/sparc-web/calc), with SMILES (simplified molecular-input line-entry system) strings as input. COSMOtherm predicts a large variety of properties based on COSMO-RS (conductor-like screening model for real solvents) theory, which uses quantum-chemical calculations and statistical thermodynamics (Klamt and Eckert, 2000;Klamt, 2005). First, TURBOMOLE (version 6.6, 2014, University of Karlsruhe & Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007 available from www.turbomole.com) optimizes the geometry of the molecules of interest at the BP-TZVP level. COSMOconf (version 3.0, COSMOlogic) then selects a maximum of ten lowest energy conformers for each calculated molecule and generates COSMO files. Calculations with TURBOMOLE and COSMOconf were performed on the General Purpose Cluster (GPC) supercomputer at the SciNet HPC Consortium at University of Toronto (Loken et al., 2010). Finally, COSMOtherm (version C30_1501 with BP_TZVP_C30_1501 parameterization, COSMOlogic GmbH & Co. KG, Leverkusen, Germany, 2015) calculates partitioning coefficients from the selected COSMO files at 15 °C.

In order to compare different predictions numerically, we calculated the mean difference (MD) and the mean absolute difference (MAD) for each pair of $K_{\mathrm{WIOM/G}}$ or $K_{\mathrm{W/G}}$ sets:

$$\mathrm{MD_{XY}} = \frac{1}{\mathrm{n}}\sum\nolimits_{\mathrm{i}} \left(\log_{10}K_{\mathrm{i,CP/G\,X}} - \log_{10}K_{\mathrm{i,CP/G\,Y}}\right) \tag{5}$$

$$\mathrm{MAD_{XY}} = \frac{1}{\mathrm{n}}\sum\nolimits_{\mathrm{i}} \left|\log_{10}K_{\mathrm{i,CP/G\,X}} - \log_{10}K_{\mathrm{i,CP/G\,Y}}\right| \tag{6}$$

where CP ("condensed phase") stands for either WIOM or water and X and Y represents two prediction techniques.

Results

The Range of Estimated Partitioning Coefficients

Partitioning coefficients predicted for each compound with different methods are given in an Excel spreadsheet as Supporting Information. All three methods predicted the log $K_{WIOM/G}$ for these organic compounds to range from approximately 0 to 15 (Figure 1 (a)-(c)). Hodzic et al. (2014) predicted a log $K_{WIOM/G}$ in the range of approximately 0 and 20 at 25 °C (see conversion between C^* and $K_{WIOM/G}$ in the supporting information) for oxidation products of different VOCs (including n-alkanes, benzene, toluene, xylene, isoprene and terpenes), i.e. their data set included higher $K_{WIOM/G}$ values (indicating generally lower volatility) than those generated here, even though $K_{WIOM/G}$ values are lower at higher temperature.

The log $K_{W/G}$ range predicted for the studied compounds by the three methods is more variable (Figure 1 (d)-(f)), with the ABSOLV/ppLFER predictions covering a wider range (-1.4 to 21.3) than either SPARC (-2.7 to 17.2) or COSMOtherm (-2 to 13.8). Hodzic et al. (2014) predicted a log $K_{W/G}$ in the range of -2.6 and 17.4 at 25°C (see conversion to K_H in the supporting information). The wider range of the ABSOLV/ppLFER predictions is due to much higher predicted $K_{W/G}$-values for compounds with the highest affinity for the aqueous phase.

Comparison between Different Prediction Methods

The discrepancies between different predictions (MAD and MD) are given in Table 1. Figure S1 in the supporting information illustrates the frequency of the discrepancies between different pairs of predicted log $K_{WIMO/G}$ and log $K_{W/G}$ values. This discrepancy only indicates the agreement between any two predictions with little indication of the accuracy of the prediction, for reasons discussed later. The agreement between the $K_{WIOM/G}$ predictions by COSMOtherm, SPARC and ABSOLV/ppLFER was reasonable (Figure 1 (a)-(c)). In particular, the MAD between $K_{WIOM/G}$ predictions is less than 1 log units (Table 1) and therefore similar to what had been previously found for a much smaller set of *n*-alkane oxidation products (Wania et al., 2014). The $K_{WIOM/G}$-values predicted by SPARC tend to be higher than those predicted by COSMOtherm and ABSOLV/ppLFER (MD of -0.64 and -0.79 in log units, respectively), whereas the latter two predictions have a slightly better agreement, with a MD of 0.15 log units (Figure 1 (c) and Table 1). Overall, the agreement in the $K_{\mathrm{WIOM/G}}$ predicted with these three methods, which are based on very different theoretical foundations, is much better than that between different vapor pressure estimation methods commonly used for gas-particle partitioning calculations (Valorso et al., 2011).

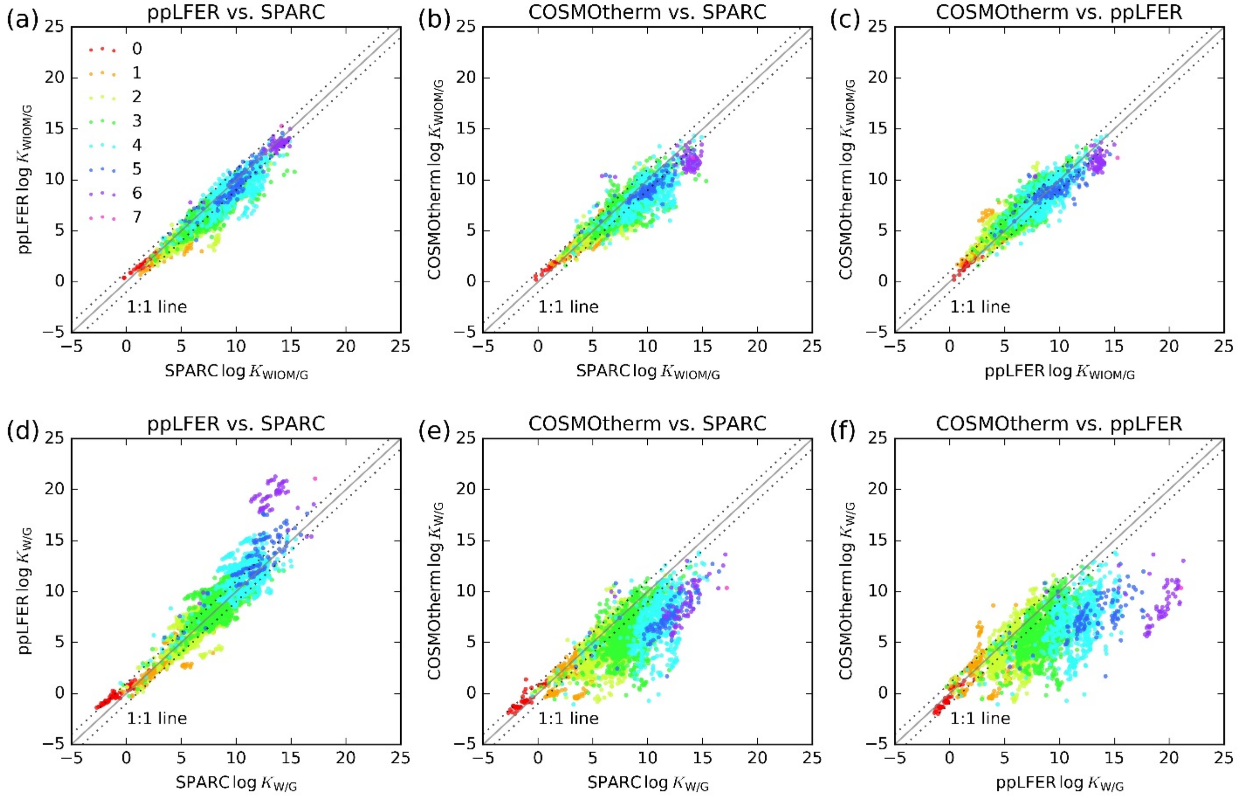


Figure 1 Comparison of the $K_{\mathrm{WIOM/G}}$ (upper panel) and $K_{\mathrm{W/G}}$ (lower panel) predicted using COSMOtherm, SPARC and ABSOLV/ppLFERs. The differently colored dots indicate the number of functional groups in the molecules. The solid line indicates a 1:1 agreement. The dotted lines indicate a deviation by ±1 log unit.

The $K_{\mathrm{W/G}}$ predicted by ABSOLV/ppLFER and SPARC differ from COSMOtherm predictions substantially, on average by more than two orders of magnitude. In Figure 1 (e) and (f), predictions are more scattered (indicating a larger MAD) and most markers are located above the 1:1 line, indicating that $K_{\mathrm{W/G}}$ predicted by COSMOtherm are mostly lower than those predicted by SPARC and ABSOLV/ppLFER, with a MD of -2.06 and -2.42 log units, respectively.

These discrepancies tend to increase with the $K_{W/G}$. Raventos-Duran et al. (2010) also showed that the reliability of $K_{W/G}$ estimates made by GROMHE, SPARC and HENRYWIN decreases with increasing affinity for the aqueous phase. $K_{W/G}$ predictions by SPARC and ABSOLV/ppLFER are more consistent (with a MAD around 1 log units, see Figure 1 (d)). The largest discrepancies between ABSOLV/ppLFER and SPARC (and also between ABSOLV/ppLFER and COSMOtherm) occur for compounds with the highest $K_{W/G}$ as predicted by ABSOLV/ppLFER (purple markers in Figure 1 (d) and (f)). Further analysis indicates that these compounds have the largest number of functional groups (≥6) and oxygen (9~12 oxygen) in the molecule; this will be discussed in detail below.

Table 1 Mean absolute differences (MAD) and mean differences (MD) between SPARC, ABSOLV/ppLFER and COSMOtherm predictions for compounds with different numbers of functional groups

Number of Functional Groups			0	1	2	3	4	5	>5	All
Number of Compounds			63	372	1179	1064	565	111	60	3414
log$K_{WIOM/G}$	ppLFER vs. SPARC	MAD	0.24	0.70	0.95	0.93	1.08	0.75	0.54	0.91
		MD	0.05	-0.70	-0.91	-0.81	-0.83	-0.24	-0.30	-0.79
	COSMOtherm vs. SPARC	MAD	0.36	0.48	0.80	0.94	1.42	1.22	2.11	0.94
		MD	0.29	-0.19	-0.55	-0.57	-1.21	-0.78	-1.83	-0.64
	COSMOtherm vs. ppLFER	MAD	0.32	0.67	0.63	0.74	0.89	0.93	1.72	0.73
		MD	0.24	0.51	0.36	0.24	-0.38	-0.54	-1.53	0.15
log$K_{W/G}$	ppLFER vs. SPARC	MAD	0.75	0.57	0.84	1.08	1.48	1.53	5.78	1.10
		MD	0.74	-0.09	-0.15	0.38	0.87	1.45	5.76	0.36
	COSMOtherm vs. SPARC	MAD	0.51	0.86	1.61	2.31	3.78	4.34	4.55	2.23
		MD	0.48	-0.59	-1.44	-2.18	-3.74	-4.04	-4.36	-2.06
	COSMOtherm vs. ppLFER	MAD	0.40	1.16	1.64	2.63	4.62	5.55	10.09	2.64
		MD	-0.26	-0.50	-1.29	-2.56	-4.61	-5.50	-10.05	-2.42

Dependence of Partitioning Coefficients on Attributes of the Compounds

The equilibrium partitioning coefficients depend on molecular attributes. Here we explored this dependency on the number of functional groups, molecular mass, generation of oxidation, number of oxygens and O:C ratio.

Previous work observed that discrepancies between vapor pressure predictions by different methods increased with the number of functional groups in atmospherically relevant organic compounds (Valorso et al., 2011;Barley and McFiggans, 2010;Compernolle et al., 2011). For instance, the MAD between different vapor pressure predictions increased from 0.47 to 3.6 log units when the number of functional groups in the molecules increased from one to more than three (Valorso et al., 2011). In order to explore if the partitioning coefficients predicted with SPARC, ABSOLV/ppLFER and COSMOtherm show the same dependence on the number of functional groups, we counted the number of hydroxyl (ROH), aldehyde (RCHO), ketone (RCOR'), carboxylic acid (RCOOH), ester (RCOOR'), ether (ROR'), peracid (RCOOOH), peroxide (ROOH, ROOR'), nitrate (NO3), peroxyacyl nitrate (PAN), nitro (NO2) groups, halogen (Cl, Br), and sulphur (S) in the 3414 molecules. About two thirds (2243) of the compounds contain two or three functional groups (Table 1). 736 compounds contain more than three functional groups and the rest contains just one or no functional group. In Figure 1 the compounds are colored according to the number of functional groups in a molecule and Table 1 lists the MAD and MD between predictions based on the number of functional groups. The predicted partitioning coefficients (both $K_{WIOM/G}$ and $K_{W/G}$) generally increase with the number of functional groups (Figure 1 and Figure S2). Compounds with no functional groups are the precursor compounds, which generally have a smaller discrepancy among different prediction methods.

The boxplots in Figure 2 show the difference in SPARC, ABSOLV/ppLFER and COSMOtherm predictions for compounds having different number of functional groups. The mean absolute difference in predicted log $K_{WIOM/G}$ is mostly (and on average) smaller than one log unit for compounds with up to seven functional groups (Table 1). There is a slightly larger discrepancy in the predicted log $K_{WIOM/G}$ values for compounds with more than three functional groups. The agreement among different methods does not deteriorate as much with increasing number of functional groups as that among vapor pressure predictions. The largest MADs of 1.72 and 2.11 between COSMOtherm and ABSOLV/ppLFER, and between COSMOtherm and SPARC, respectively, for compounds with >5 functional groups (Table 1) are still much lower than discrepancies reported between different vapor pressure prediction methods (Valorso et al., 2011).

Different from the predictions for $K_{\mathrm{WIOM/G}}$, the discrepancy between COSMOtherm and SPARC and between COSMOtherm and ABSOLV/ppLFER in the predicted $K_{\mathrm{W/G}}$ increases significantly with the number of functional groups (Figures 1 and 2), from less than one order of magnitude for compounds with no functional groups to up to five orders of magnitude for compounds with more than three functional groups (Table 1). In addition, the MD in Table 1 and Figure 2 indicate that the discrepancies are almost always in one specific direction, i.e. a lower value of $K_{\mathrm{W/G}}$ estimated by COSMOtherm. This is evidenced by the almost identical absolute values of MAD and MD between COSMOtherm and ABSOLV/ppLFER and between COSMOtherm and SPARC for compounds with more than three functional groups (Table 1). The uncertainty of the SPARC, ABSOLV/ppLFER and COSMOtherm predictions of $K_{\mathrm{W/G}}$ tends to increase with the number of functional groups. Clearly, the reliability of $K_{\mathrm{W/G}}$ estimates for multifunctional compounds needs further assessment.

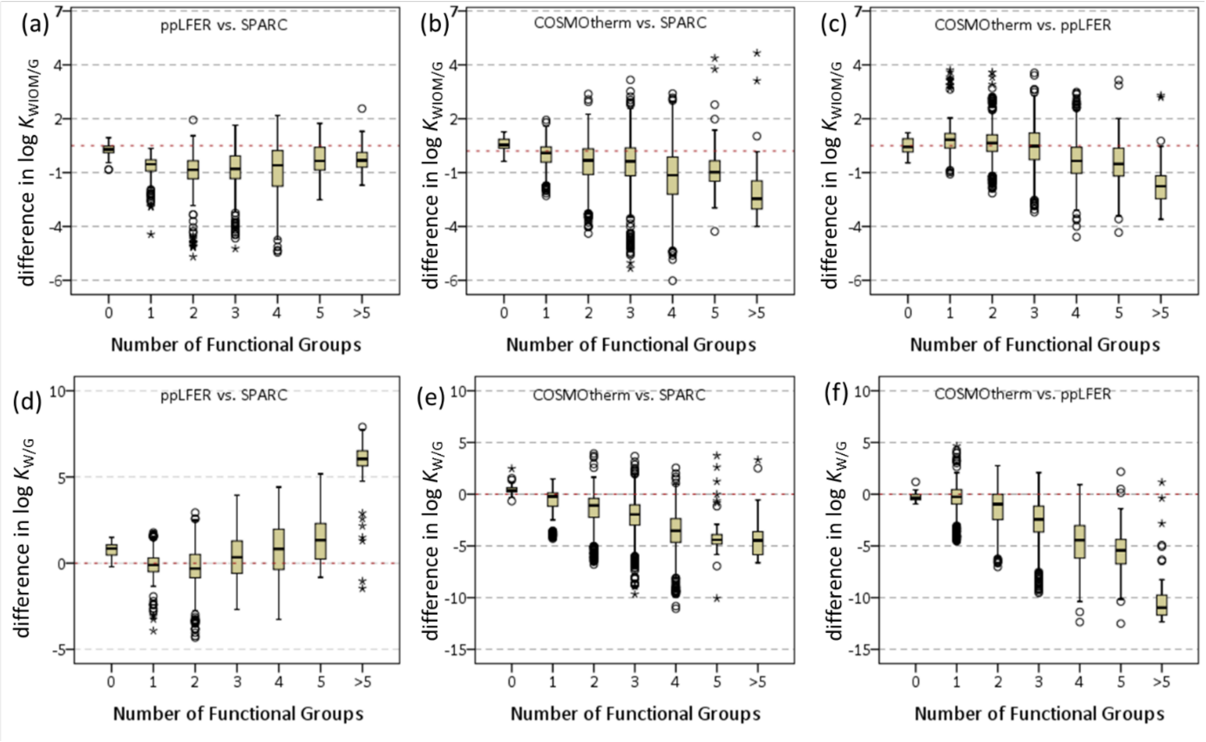


Figure 2 Boxplot of difference in SPARC, ABSOLV/ppLFER and COSMOtherm predictions for compounds with different number of functional groups. The line inside each box shows the median difference for log $K_{\mathrm{WIOM/G}}$ or log $K_{\mathrm{W/G}}$ for different categories of compounds.

The marker circle and star indicates possible outliers and extreme values, respectively. Note the different scales for different panels.

It is also possible to explore the dependence of the prediction discrepancy on other molecular attributes, such as molecular mass (Figures S3 and S4), the number of oxygen in the molecule (Figures S5 and S6), the O:C ratio (Figure S7), the number of oxidation steps a molecular has undergone (oxidation generation, Figure S8), or the number of occurrences of a specific type of functional group, e.g. hydroxyl, in a molecule (Figure S9). The prediction discrepancies become larger with an increase in each of these parameters, especially for $K_{\mathrm{W/G}}$. This is not surprising as these molecular attributes all tend to be highly correlated, i.e. with each oxidation step a molecule becomes more oxygenated, has a large molar mass, a larger number of oxygen, a higher O:C ratio, and a larger number of functional groups.

Discussion

We believe there are primarily two factors that are contributing to errors in the prediction of $K_{\mathrm{CP/G}}$ for the SOA compounds. One is the lack of experimental data for compounds that are similar to the SOA compounds, which implies that prediction methods relying on calibration with experimental data are being used outside their applicability domain. The other is the failure of some prediction methods to account for the various conformations that compounds with multiple functional groups can undergo due to extensive intra-molecular interaction (mostly internal hydrogen bonding, see Figure S10 for example). The two factors are related: in some instances a prediction method cannot account for such conformations precisely because the calibration data set does not contain compounds that undergo such intra-molecular interactions.

SPARC relies to some extent on calibrations with empirical data. While the experimental data underlying SPARC have not been disclosed, it is highly unlikely that they include multifunctional compounds of atmospheric relevance (e.g. compounds containing multiple functional groups, including peroxides, peroxy acids etc.), simply because such empirical data do not exist. It is therefore safe to assume that many of the 3414 SOA compounds will fall outside of the domain of applicability of SPARC. It is also likely that SPARC can only account for intra-molecular interactions and conformations to a limited extent, if at all.

In the case of ppLFER, there are actually two predictions that rely on calibration with empirical data, the prediction of solute descriptors and the prediction of $K_{\mathrm{CP/G}}$. The solute descriptors are predicted with ABSOLV, because experimentally measured descriptors are unavailable for multifunctional atmospheric oxidation products. ABSOLV relies on a group contribution approach (Platts et al., 1999) complemented by some other, undisclosed procedures that make use of experimental partitioning coefficients between various phases (ACD/Labs, 2016). Again, those experimental data do not comprise compounds structurally similar to the multifunctional atmospheric oxidation products considered here. As a group contribution method, which adds up the contributions of different functional groups to a compound's property, ABSOLV therefore cannot, or only to a limited extent, consider the interactions between different functional groups in a molecule.

Ideally, when supplied with well-characterized solute descriptors, ppLFERs should be able to consider the influence of both intra-molecular interactions and the interactions a molecule has with its surroundings, i.e. the involved partitioning phases. Even if a molecule has different conformations in different phases, i.e. if the solute descriptors for a compound are phase dependent, it is possible to derive well-calibrated "average" descriptors to use in a ppLFER (Niederer and Goss, 2008). However ABSOLV cannot correctly predict such "average" descriptors and our ppLFER predictions therefore cannot account for the influence of conformations.

In the case of the actual ppLFER prediction of $K_{\mathrm{W/G}}$ and $K_{\mathrm{WIOM/G}}$, the empirical calibration datasets are public (Goss, 2006;Arp et al., 2008) and do not comprise compounds that are representative of the 3414 SOA compounds in terms of the number of functional groups per molecule or the range of K-values. For instance, the log $K_{\mathrm{W/G}}$ of the 217 compounds Goss (2006) used for the development of a ppLFER ranged from -2.4 to 7.4, i.e. the highest $K_{\mathrm{W/G}}$ predicted here is almost 14 orders of magnitude higher than the highest $K_{\mathrm{W/G}}$ included in the calibration. Similarly, Arp and Goss (2009) developed the ppLFERs for atmospheric aerosol from an empirical dataset of 50~59 chemicals, whose log $K_{WIOM/G}$ ranged from approximately 2 to 7. The highest $K_{WIOM/G}$ predicted here is eight orders of magnitude higher. Predictions for compounds outside of the calibration domain may introduce large errors and the high $K_{W/G}$ and $K_{WIOM/G}$ values estimated by ppLFER can thus be expected to be highly uncertain. Overall, however, we expect the uncertainty of the ABSOLV-predicted solute descriptors to be larger than the uncertainty introduced by the ppLFER equation, especially for the relatively well-calibrated water/gas phase partition system. While the use of measured solute descriptors therefore would likely greatly improve the ppLFER prediction (Endo and Goss, 2014), those are unlikely to become available for atmospheric oxidation products.

In contrast to the other methods, COSMOtherm relies only in a very fundamental way on some empirical calibrations (and these calibrations are not specific for specific compound classes or partition systems) and it considers intra-molecular interactions and the different conformations of a molecule. As such, COSMOtherm is not constrained by the limitations the other methods face, namely the lack of suitable calibration data, which necessitates extreme extrapolations and predictions beyond the applicability domain, and the failure to account for the effect of intra-molecular interactions and conformations on the interactions with condensed phases.

Because intra-molecular interactions are likely to reduce the potential of a compound to interact with condensed phases (i.e. the organic and aqueous phase), ignoring them can be expected to lead to overestimated partitioning coefficients $K_{CP/G}$ and to underestimated vapor pressures (P_L) and C^*, i.e. underestimating the volatility of the organic compounds. This is consistent with COSMOtherm-predicted $K_{WIOM/G}$ and $K_{W/G}$-values for multifunctional compounds that are lower than the SPARC and ABSOLV/ppLFER predictions (i.e. MD<0 in Table 1), because the latter do not account for the influence of intra-molecular interactions. Kurtén et al. (2016) similarly found that COSMOtherm-predicted saturation vapor pressures for most of the more highly oxidized monomers were significantly higher (up to 8 orders of magnitude) than those predicted by group-contribution methods. The wider range on the higher end of the log C^* values estimated by Hodzic et al. (2014) is possibly due to the large uncertainties associated with vapor pressure estimation (likely underestimation) for low volatile compounds. Valorso et al. (2011) also found group contribution methods to underestimate the saturation vapor pressure of multifunctional species.

Compared to $K_{WIOM/G}$, P_L and C^*, ignoring intra-molecular interaction is likely even more problematic in the case of $K_{W/G}$ prediction. Intra-molecular interactions mostly affect the ability of the molecule to undergo H-bonding with solvent molecules. The system constants describing H-bond interactions (a and b) are larger in the ppLFER equations for $K_{W/G}$ than in the one for $K_{WIOM/G}$ (Arp et al., 2008;Goss, 2006), indicating a stronger effect of H-bonds on water/gas partitioning than WIOM/gas partitioning. This likely is the reason why the COSMOtherm-predicted $K_{W/G}$ are so much lower than the $K_{W/G}$ predicted by the other two methods, whereas the difference is much smaller for the $K_{WIOM/G}$ (Table 1). It likely also explains why the discrepancies among the predicted $K_{W/G}$ increase with the number of functional groups. It is more difficult to predict $K_{W/G}$ than $K_{WIOM/G}$, because the free energy cost of cavity formation in water is influenced more strongly by H-bonding and therefore much more variable than in WIOM. Certainly, the activity coefficient in water (γ_W) is much more variable than the activity coefficient in WIOM (γ_{WIOM}) for the investigated substances. log γ_{WIOM} predicted by COSMOtherm at 15 °C varies from -3.8 to 1.8 (with an average of 0.04, indicating a γ_{WIOM} close to unity, and a standard deviation of 0.5, 94 % of the compounds have a log γ_{WIOM} between -1 and 1), whereas γ_W ranges from -2.3 to 8.9 (with an average of 2.7 and a standard deviation of 1.4) (Supporting information Excel spreadsheet and Figure S11).

In the absence of experimental data for multi-functional SOA compounds, we do not know whether COSMOtherm-predicted $K_{W/G}$ and $K_{WIOM/G}$ values are any better than the other predictions. For example, two earlier studies suggested that COSMOtherm might be overestimating vapor pressures of multi-functional oxygen-containing compounds (Kurtén et al., 2016;Schröder et al., 2016). However, we can infer that:

- the fact that COSMOtherm on the one hand and ABSOLV/ppLFERs and SPARC on the other hand predict $K_{WIOM/G}$ that are on average within one order of magnitude for all studied compounds, and less than two orders of magnitude for highly oxygenated multifunctional organic compounds, lends credibility to all three predictions and suggests that partly ignoring intra-molecular interactions and extrapolating beyond the applicability domain incurs only limited errors in the $K_{WIOM/G}$ prediction of ABSOLV/ppLFERs and SPARC. In addition, COSMOtherm and SPARC use a single surrogate molecule to represent the WIOM phase, while ppLFERs were calibrated from atmospheric aerosols. The agreement among different methods suggests that the surrogate suitably represents the solvation properties of organic aerosol.

- the generally better agreement between $K_{W/G}$ values predicted by ABSOLV/ppLFER and SPARC (Figure 1(d)) should not be seen as an indication that these methods are better at predicting $K_{W/G}$. In fact, the lower $K_{W/G}$ values predicted by COSMOtherm have a higher chance of being correct than the $K_{W/G}$ values predicted by ABSOLV/ppLFER and SPARC.

While ABSOLV/ppLFERs, SPARC and the group contributions methods currently used in the atmospheric chemistry community are much more easily implemented for the large number of compounds implicated in SOA formation, the current study demonstrates that the expertise and time required to perform quantum-chemical calculations for atmospherically relevant molecules should constitute but a minor impediment to a wider adoption of COSMOtherm predictions. Here, we are not only compiling all the predictions we have made in the supporting information file, we are also making available the cosmo-files (see Data Availability for details), whose generation is the major time and CPU-demanding step in the use of COSMOtherm.

Atmospheric Implications

The phase distribution of an organic compound in the atmosphere depends on its partitioning coefficients. The two-dimensional partitioning space defined by log $K_{W/G}$ and log $K_{WIOM/G}$ introduced recently (Wania et al., 2015) is used here to illustrate the difference in the equilibrium phase distribution of these compounds in the atmosphere that arises from using partitioning coefficients estimated by different methods (Figure 3). A detailed description of partitioning space has been provided by Wania et al. (2015), a brief explanation is given in the supporting information (Figure S12). Briefly, the blue solid lines between the differently colored fields indicate partitioning property combinations that lead to equal distributions between two phases in a phase-separated aerosol scenario, with a liquid water content (LWC) of 10 µg/m^3

and organic matter loading (OM) of 10 µg/m^3. The blue dotted lines represent a cloud scenario where LWC is 0.3 g/m^3 and OM is 10 µg/m^3. Figures S13 and S14 in the supporting information show an aerosol scenario without an aqueous phase and a cloud scenario without a separated organic phase because all of the OM is dissolved in the aqueous phase (see also Figure S12 (c) and (d)). Compounds are located in the partitioning space based on their estimated partitioning coefficients ($K_{WIOM/G}$ and $K_{W/G}$). Compounds on the boundary lines have 50 % in either of the two phases on both sides of the boundary and are thus most sensitive to uncertain partitioning properties. On the other hand, for substances that fall far from the boundary lines indicating a phase transition (e.g. volatile compounds with two or less functional groups), even relatively large uncertainties in the partitioning coefficients could be tolerated, because they are inconsequential.

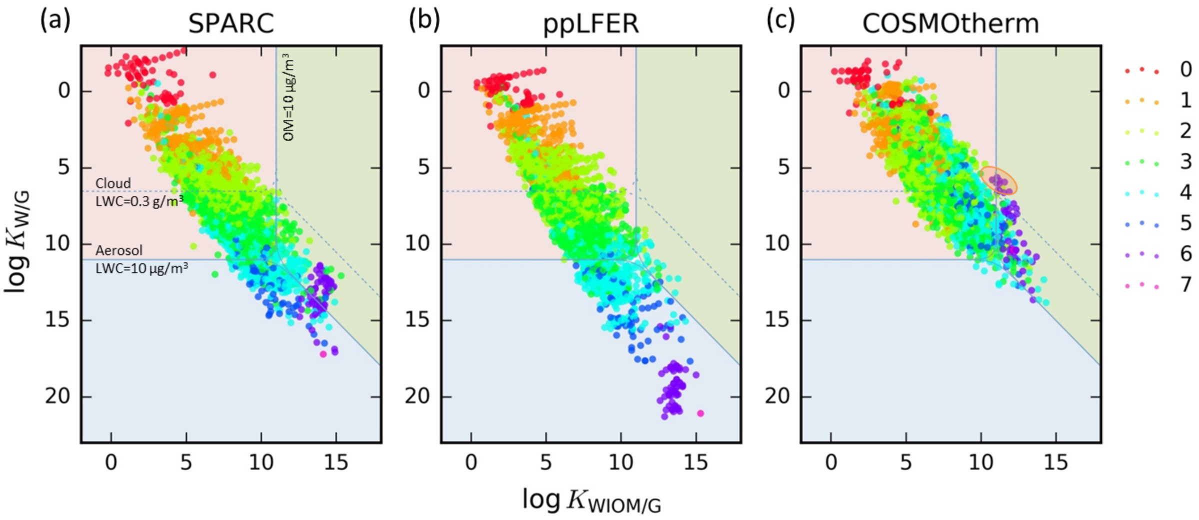


Figure 3 Partitioning space plot, showing in pink, blue and green the combinations of partitioning properties that lead to dominant equilibrium partitioning to the gas, aqueous, and WIOM phases, respectively. The blue solid and dotted lines are boundaries for an aerosol scenario (LWC 10 µg/m^3, 10 µg/m^3 OM) and a cloud scenario (LWC 0.3 g/m^3, 10 µg/m^3 OM), respectively. The differently colored dots indicate the number of functional groups in the molecules.

When plotted in the chemical partition space, the 3414 chemicals occupy more or less the same region as the much smaller set of SOA compounds investigated earlier (Wania et al., 2015). When using predictions by COSMOtherm the SOA compounds cover a relatively smaller region as compared to ABSOLV/ppLFER and SPARC. With increasing number of functional groups (Figure 3) or molecular weight (Figure S15), an increasing fraction of these compounds partitions into the condensed phases, i.e. WIOM or water. In general, compounds with water or WIOM as the dominant phase usually are multifunctional, i.e. contain more than two functional groups. According to Figure S15, compounds with predominant partitioning into WIOM usually have a molar mass in excess of 200 g/mol, while some compounds with molar mass less than 200 g/mol prefer the aqueous phase. Other than the water content and WIOM loadings illustrated in Figure 3, in reality a compound's atmospheric phase distribution depends on other factors such as the organic matter composition, salt content, pH, and temperature (Wania et al., 2015;Wang et al., 2015).

Comparing the different panels of Figure 3 reveals that the atmospheric equilibrium phase distribution of SOA compounds can be very different depending on which methods is used for partitioning coefficient estimation. The difference is most striking when comparing the placement of highly functionalized compounds (with more than 3 functional groups) based on ABSOLV/ppLFER and COSMOtherm predictions. The large $K_{W/G}$ values estimated by ABSOLV/ppLFERs lead to these compounds having a high affinity for aqueous aerosol. In contrast, predictions by COSMOtherm suggest that only very few of them (and not even the ones with the highest number of functional groups) prefer the aqueous aerosol phase; instead most of them have either gas or WIOM as the dominant phase. SPARC predicts a slightly larger preference of highly functionalized compounds for the aqueous phase than COSMOtherm.

In a cloud scenario with a much higher LWC (shown by the blue dotted boundary lines in Figure 3), the choice of $K_{W/G}$ prediction method also matters. Whereas with ABSOLV/ppLFER and SPARC most of the highly functionalized compounds (i.e. 96 % or 97 % of the 736 compounds with >3 functional groups) partitions into aqueous phase, only two-thirds (64 %) do so when the $K_{W/G}$s predicted by COSMOtherm are used. Further, only COSMOtherm predicts that some of the SOA compounds (circled in Figure 3 (c)) would prefer to form a separate WIOM phase rather than dissolve in the bulk aqueous phase. Those compounds are not sufficiently soluble in water to partition to the cloud and are not sufficiently volatile to be in the gas phase.

Table 2 summarizes the number and percentage of compounds that have dominant partitioning (at least 50 %) into different phases, which shows the impact of using different prediction techniques on phase distribution calculations in different atmospheric scenarios. In a parameterisation of SOA formation that includes an aqueous aerosol phase, use of $K_{W/G}$ predicted by ABSOLV/ppLFERs (and probably also the commonly employed group contribution methods) would lead to much higher SOA mass than use of $K_{W/G}$ predicted by COSMOtherm. For instance, 10 % and 17 % of the compounds predominantly partition into the aqueous phase when predictions by SPARC and ABSOLV/ppLFER are used, in contrast to only 14 compounds (less than 1 %) with COSMOtherm predictions (Table 2 scenario (a)). A large difference also occurs in the cloud scenarios (Table 2 scenarios (b) and (d)), where SPARC and ABSOLV/ppLFER predict twice as many compounds partitioning into the aqueous phase than COSMOtherm. Incidentally, in a parameterization of SOA formation that does not account for an aqueous aerosol phase (the scenario in Figure S12 (c) and Table 2 (c)), the impact of the choice of partitioning prediction method is much smaller. The number of compounds on the right side of the blue dotted boundary in Figure S13 does not vary substantially with different predictions. Table S1 in the supporting information summarizes the number and percentage of compounds that change their partitioning between gas and condensed phase under different atmospheric conditions when a different prediction method is used. Depending on the scenarios, a total of 2.0 % up to 34 % of the 3414 compounds have a different dominant phase when using a different prediction method. This change is larger for the cloud scenarios and much lower for the aerosol scenarios especially if the aerosol contains no water.

Table 2 Percentage and number of compounds with at least 50 % in gas, water or WIOM phase under different aerosol and cloud scenarios predicted with SPARC, ABSOLV/ppLFER and COSMOtherm. The four scenarios (a-d) correspond to the scenarios in Figure S12 (a-d) in Supporting information.

aerosol scenarios	(a) (LWC=10 μg/m^3, OM=10 μg/m^3)			(c) without water phase (LWC=0 μg/m^3, OM=10 μg/m^3)	
	Φ_G>50 %[a]	Φ_W>50 %[a]	Φ_{WIOM}>50 %[a]	Φ_G>50 %	Φ_{WIOM}>50 %
SPARC	85 % (2892)[b]	10 % (352)	4% (134)	92 % (3132)	8 % (282)
ABSOLV/ppLFER	82 % (2804)	17 % (570)	1% (25)	96 % (3267)	4 % (141)
COSMOtherm	96 % (3268)	0 % (14)	3% (119)	96 % (3282)	4 % (131)
cloud scenarios	(b) (LWC=0.3 g/m^3, OM=10 μg/m^3)			(d) without WIOM phase (LWC=0.3 g/m^3, OM=0 μg/m^3)	
	Φ_G>50%	Φ_W>50 %	Φ_{WIOM}>50 %	Φ_G>50 %	Φ_W>50 %
SPARC	36 % (1242)	64 % (2168)	0 % (0)	36 % (1242)	64 % (2172)
ABSOLV/ppLFER	35 % (1201)	65 % (2211)	0 % (0)	35 % (1203)	65 % (2211)
COSMOtherm	66 % (2258)	33 % (1137)	0 % (9)	66 % (2267)	34 % (1147)

[a] Φ_G, Φ_W and Φ_{WIOM} represent for fractions of compounds in gas phase, water phase and WIOM phase, respectively.

[b] number in brackets are number of compounds

Conclusions

For compounds implicated in SOA formation, the prediction of $K_{W/G}$ is much more uncertain than the prediction of $K_{WIOM/G}$. This is true even if we consider that $K_{WIOM/G}$ will vary somewhat depending on the composition of the WIOM (Wang et al., 2015). In particular, the methods currently used for $K_{W/G}$ prediction of these substances have the potential to greatly overestimate $K_{W/G}$. This uncertainty is consequential, as the predicted equilibrium phase distribution in the atmosphere, and therefore also the predicted aerosol yield, is very sensitive to the predicted values of $K_{W/G}$: depending on the method used for prediction, the aqueous phase is either very important for SOA formation from the studied set of compounds or hardly at all. Isaacman-VanWertz et al. (2016) recently found the estimated phase distribution of 2-methylerythritol, an isoprene oxidation product (in Figure S6), highly dependent on the chosen method for predicting $K_{W/G}$. Here we show that this is a general issue potentially affecting a very large number of SOA compounds. In order to identify reliable prediction methods, it will be necessary to experimentally determine the phase distribution of highly functionalized, atmospherically relevant substances, whereby the focus should be on establishing their partitioning into aqueous aerosol.

Data Availability

COSMO files for the 3414 organic compounds can be accessed by contacting the corresponding author.

Supporting Information

The supporting information contains figures and text mentioned in the paper, including detailed information on the organic compounds, e.g. SMILES, molecular formula, molecular weight, functional groups, O:C ratios, predicted *K*-values, ABSOLV predicted solute descriptors, COSMOtherm predicted vapor pressures and activity coefficients in WIOM and water.

Acknowledgement

We acknowledge funding from Natural Sciences and Engineering Research Council of Canada. Computations for the COSMO files in this study were performed on the General Purpose Cluster (GPC) supercomputer at the SciNet HPC Consortium. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto.

References

ACD/Labs: Advanced Chemistry Development. Data sheet: Absolv prediction module. Toronto (Canada). [cited 2016 December 2]. Available from: http://www.acdlabs.com/download/docs/datasheets/datasheet_absolv.pdf., 2016.

Arp, H. P. H., Schwarzenbach, R. P., and Goss, K. U.: Ambient gas/particle partitioning. 2: The influence of particle source and temperature on sorption to dry terrestrial aerosols, Environmental Science & Technology, 42, 5951-5957, 10.1021/es703096p, 2008.

Arp, H. P. H., and Goss, K. U.: Ambient Gas/Particle Partitioning. 3. Estimating Partition Coefficients of Apolar, Polar, and Ionizable Organic Compounds by Their Molecular Structure, Environmental Science & Technology, 43, 1923-1929, 10.1021/es8025165, 2009.

Barley, M. H., and McFiggans, G.: The critical assessment of vapour pressure estimation methods for use in modelling the formation of atmospheric organic aerosol, Atmospheric Chemistry and Physics, 10, 749-767, 10.5194/acp-10-749-2010, 2010.

Compernolle, S., Ceulemans, K., and Muller, J. F.: EVAPORATION: a new vapour pressure estimation method for organic molecules including non-additivity and intramolecular interactions, Atmospheric Chemistry and Physics, 11, 9431-9450, 10.5194/acp-11-9431-2011, 2011.

Endo, S., Pfennigsdorff, A., and Goss, K. U.: Salting-out effect in aqueous NaCl solutions: trends with size and polarity of solute molecules, Environmental Science & Technology, 46, 1496-1503, 10.1021/es203183z, 2012.

Endo, S., Goss, K.-U., Applications of polyparameter linear free energy relationships in environmental chemistry, Environmental Science and Technology, 48, 12477-12491, 2014.

Goss, K.-U.: Prediction of the Temperature Dependency of Henry's Law Constant using Poly-Parameter Linear Free Energy Relationships, Chemosphere, 64, 1369-1374, 2006.

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prevot, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, Atmospheric Chemistry and Physics, 9, 5155-5236, 2009.

Hilal, S. H., Karickhoff, S. W., and Carreira, L. A.: Prediction of the solubility, activity coefficient and liquid/liquid partition coefficient of organic compounds, Qsar & Combinatorial Science, 23, 709-720, 2004.

Hilal, S. H., Ayyampalayam, S. N., and Carreira, L. A.: Air-Liquid Partition Coefficient for a Diverse Set of Organic Compounds: Henry`s Law Constant in Water and Hexadecane, Environmental Science & Technology, 42, 9231-9236, doi:10.1021/es8005783, 2008.

Hodzic, A., Aumont, B., Knote, C., Lee-Taylor, J., Madronich, S., and Tyndall, G.: Volatility dependence of Henry's law constants of condensable organics: Application to estimate depositional loss of secondary organic aerosols, Geophysical Research Letters, 41, 4795-4804, 10.1002/2014gl060649, 2014.

Ip, H. S. S., Huang, X. H. H., and Yu, J. Z.: Effective Henry's law constants of glyoxal, glyoxylic acid, and glycolic acid, Geophysical Research Letters, 36, 10.1029/2008gl036212, 2009.

Isaacman-VanWertz, G., Yee, L. D., Kreisberg, N. M., Wernis, R., Moss, J. A., Hering, S. V., de Sá, S. S., Martin, S. T., Alexander, M. L., Palm, B. B., Hu, W., Campuzano-Jost, P., Day, D. A., Jimenez, J. L., Riva, M., Surratt, J. D., Viegas, J., Manzi, A., Edgerton, E., Baumann, K., Souza, R., Artaxo, P., and Goldstein, A. H.: Ambient Gas-Particle Partitioning of Tracers for Biogenic Oxidation, Environmental Science & Technology, 50, 9952-9962, 10.1021/acs.est.6b01674, 2016.

Kalberer, M., Paulsen, D., Sax, M., Steinbacher, M., Dommen, J., Prevot, A. S. H., Fisseha, R., Weingartner, E., Frankevich, V., Zenobi, R., and Baltensperger, U.: Identification of polymers as major components of atmospheric organic aerosols, Science, 303, 1659-1662, 10.1126/science.1092185, 2004.

Klamt, A., and Eckert, F.: COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids, Fluid Phase Equilibria, 172, 43-72, 2000.

Klamt, A.: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design, Elsevier, Amsterdam, 2005.

Kroll, J. H., Donahue, N. M., Jimenez, J. L., Kessler, S. H., Canagaratna, M. R., Wilson, K. R., Altieri, K. E., Mazzoleni, L. R., Wozniak, A. S., Bluhm, H., Mysak, E. R., Smith, J. D., Kolb, C. E., and Worsnop, D. R.: Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol, Nature Chemistry, 3, 133-139, http://www.nature.com/nchem/journal/v3/n2/abs/nchem.948.html#supplementary-information, 2011.

Kurtén, T., Tiusanen, K., Roldin, P., Rissanen, M., Luy, J.-N., Boy, M., Ehn, M., and Donahue, N.: α-Pinene autoxidation products may not have extremely low saturation vapor pressures despite high O:C ratios, The Journal of Physical Chemistry A, 120, 2569-2582, 10.1021/acs.jpca.6b02196, 2016.

Loken, C., Gruner, D., Groer, L., Peltier, R., Bunn, N., Craig, M., Henriques, T., Dempsey, J., Yu, C.-H., Chen, J., Dursi, L. J., Chong, J., Northrup, S., Pinto, J., Knecht, N., and Van Zon, R.: SciNet: Lessons Learned from Building a Power-efficient Top-20 System and Data Centre, Journal of Physics: Conference Series, 256, 012026, 2010.

McFiggans, G., Topping, D. O., and Barley, M. H.: The sensitivity of secondary organic aerosol component partitioning to the predictions of component properties - Part 1: A systematic evaluation of some available estimation techniques, Atmospheric Chemistry and Physics, 10, 10255-10272, 10.5194/acp-10-10255-2010, 2010.

Mouchel-Vallon, C., Bräuer, P., Camredon, M., Valorso, R., Madronich, S., Herrmann, H., and Aumont, B.: Explicit modeling of volatile organic compounds partitioning in the atmospheric aqueous phase, Atmos. Chem. Phys., 13, 1023-1037, 10.5194/acp-13-1023-2013, 2013.

Niederer, C., and Goss, K.-U.: Effect of ortho-chlorine substitution on the partition behavior of chlorophenols, Chemosphere, 71, 697-702, 2008.

Platts, J. A., Butina, D., Abraham, M. H., and Hersey, A.: Estimation of Molecular Linear Free Energy Relation Descriptors Using a Group Contribution Approach, J. Chem. Inf. Comput. Sci., 39, 835-845, 1999.

Raventos-Duran, T., Camredon, M., Valorso, R., Mouchel-Vallon, C., and Aumont, B.: Structure-activity relationships to estimate the effective Henry's law constants of organics of atmospheric interest, Atmospheric Chemistry and Physics, 10, 7643-7654, 10.5194/acp-10-7643-2010, 2010.

Sander, R.: Compilation of Henry's law constants (version 4.0) for water as solvent, Atmospheric Chemistry and Physics, 15, 4399-4981, 10.5194/acp-15-4399-2015, 2015.

Schröder, B., Fulem, M., and Martins, M. A. R.: Vapor pressure predictions of multi-functional oxygen-containing organic compounds with COSMO-RS, Atmospheric Environment, 133, 135-144, http://dx.doi.org/10.1016/j.atmosenv.2016.03.036, 2016.

US EPA: Estimation Programs Interface Suite™ for Microsoft® Windows, v 4.11 or insert version used]. United States Environmental Protection Agency, Washington, DC, USA., 2012.

Valorso, R., Aumont, B., Camredon, M., Raventos-Duran, T., Mouchel-Vallon, C., Ng, N. L., Seinfeld, J. H., Lee-Taylor, J., and Madronich, S.: Explicit modelling of SOA formation from alpha-pinene photooxidation: sensitivity to vapour pressure estimation, Atmospheric Chemistry and Physics, 11, 6895-6910, 10.5194/acp-11-6895-2011, 2011.

Wang, C., Lei, Y. D., Endo, S., and Wania, F.: Measuring and modeling the salting-out effect in ammonium sulfate solutions, Environmental Science & Technology, 48, 13238-13245, 10.1021/es5035602, 2014.

Wang, C., Goss, K.-U., Lei, Y. D., Abbatt, J. P. D., and Wania, F.: Calculating equilibrium phase distribution during the formation of secondary organic aerosol using COSMOtherm, Environmental Science & Technology, 49, 8585-8594, 10.1021/acs.est.5b01584, 2015.

Wang, C., Lei, Y. D., and Wania, F.: Effect of Sodium Sulfate, Ammonium Chloride, Ammonium Nitrate, and Salt Mixtures on Aqueous Phase Partitioning of Organic Compounds, Environmental Science & Technology, 50, 12742-12749, 10.1021/acs.est.6b03525, 2016.

Wania, F., Lei, Y. D., Wang, C., Abbatt, J. P. D., and Goss, K. U.: Novel methods for predicting gas–particle partitioning during the formation of secondary organic aerosol, Atmospheric Chemistry and Physics, 14, 13189-13204, 10.5194/acp-14-13189-2014, 2014.

Wania, F., Lei, Y. D., Wang, C., Abbatt, J. P. D., and Goss, K. U.: Using the chemical equilibrium partitioning space to explore factors influencing the phase distribution of compounds involved in secondary organic aerosol formation, Atmospheric Chemistry and Physics, 15, 3395-3412, 10.5194/acp-15-3395-2015, 2015.

Waxman, E. M., Elm, J., Kurtén, T., Mikkelsen, K. V., Ziemann, P. J., and Volkamer, R.: Glyoxal and methylglyoxal setschenow salting constants in sulfate, nitrate, and chloride solutions: measurements and Gibbs energies, Environmental Science & Technology, 49, 11500-11508, 10.1021/acs.est.5b02782, 2015.