Review Pye et al., 2016

This study uses the regional chemical transport model CMAQ to predict the organic aerosol loading and water uptake over the eastern United States and compares the model simulations with observations from monitoring networks. A parameterization of the water uptake by the organic fraction of the aerosol particles is introduced, which should account for the positive feedback of water uptake triggering additional partitioning of semivolatile species to the particulate matter. Simulations relying on different parameterization of water uptake are compared with each other and with observations. It is found that the loading and water uptake of organic aerosol is sensitive to aerosol water interactions with semivolatile organic aerosol species. From too low to too high aerosol mass is predicted depending on how the water uptake is parameterized.

The subject of this study is timely and important. Knowledge of the physical properties of the organic aerosol is increasing, rendering a more physical representation in models feasible. While this paper goes into the right direction, there are inadequacies in the way the water uptake is parameterized. Most importantly, the water uptake in the case of liquid-liquid phase separation (LLPS) is unrealistic because it assumes that in the absence of LLPS the water associated with the inorganic ions is available for partitioning of organic species, while the water uptake of the organics themselves and solution non-ideality are neglected. However, LLPS results from solution non-ideality. The assumptions for the simulation LLPS are therefore inconsistent and insinuate a wrong perception of LLPS, which might lead to confusion. While it makes sense to use the formula by You et al. (2013) to estimate how frequently LLPS in the eastern US occurs (as shown in Fig. 10), the water uptake in the presence of LLPS cannot be calculated the way it is done in the manuscript. This simulation should therefore be removed. Zuend and Seinfeld (2012) have shown that in the case of LLPS, it suffices to assume that the phase separation into an organic and inorganic phase is complete, however, when organics and inorganic constituents are present in the same phase, activity coefficients accounting for interactions between organic and inorganic species should not be neglected as is done in all simulations presented here. Nevertheless, this study goes into the right direction and highlights the need to include solution non-idealities for further model improvements. I therefore recommend this paper for publication in ACP after the following comments have been addressed.

General comment:

It is difficult to find orientation in the manuscript. Some information is scattered, some is incomplete and some seems to be missing. Many different parameters are used. Their meanings are only given once when they are introduced and it is difficult to find this position again. A table listing all abbreviations would therefore be helpful. Also, some parameters and expressions are not used consistently throughout the manuscript (see specific comments). There is quite a bit of supporting information provided, but only partly referenced in the main text, which makes it hard to make use of it. The text in the figure captions is difficult to follow and needs to be improved.

Specific comments:

Page 6, line 12 and 13: the exponent of the equation does not seem to be correct: should it be "63,000(298K-T)/298K"? Please correct.

Page 6, line 18: was the accuracy for Henry's law coefficients determined for dilute solutions or for concentrated solutions as present in aerosol particles? The authors should comment on the accuracy of Henry's law coefficients applied to concentrated solutions using equation (20) and compare it to the accuracy reached by direct use of activity coefficients given by $\gamma$.

Page 8, equation (9): how is $D_{core}$ determined?

Page 9 line 2: what is meant by “mole-based”? Activity coefficients of organics are usually mole fraction based.

Page 9, line 15 and 16: “Inclusion of water, and even inorganic constituents, in the absorbing phase has been encouraged for simplified models in order to reproduce more detailed calculations (Zuend et al., 2010).” Yes, but at the same time also solution non-ideality has to be included, otherwise water uptake is too high. Zuend and Seinfeld (2012) state in their conclusions: “However, assuming ideality at higher RH (>60 %), will very likely lead to significant overprediction of SOA mass and total PM mass.” This is the case for organic mixtures and even more for mixtures including inorganic salts. Therefore, this sentence cannot be used as a justification for the model assumptions.

Page 9, line 27: it should be “modified Raoult’s law” (as in Zuend et al., 2010, and Zuend and Seinfeld, 2012) because, strictly, “Raoult’s law” does not include the activity coefficient $\gamma$.

Page 10, equation (20): the one-constant Margules equation is applicable to molecules of the same molecular size but different polarity. The molecular size of organic species and water is very different. Have you validated the applicability of this equation to aqueous organic solutions?

Page 11, line 25: Figure 3 should be explained better, here or in the figure caption. The meaning of the dashed arrow should be given explicitly.

Page 11, line 27: as explained in the general comment, the assumptions of the simulation LLPS are unphysical and inconsistent. Therefore, this simulation should be removed. It does not represent the water uptake of a system with LLPS.

Page 11, line 30: The assumptions of the second simulation should be explained better by referring to the relevant equations.

Page 12, line 3: The meaning of “a posteriori parameters” should be explained better. How were they determined?

Page 12, line 9 and 10: “The total nonvolatile POA in CMAQ is assumed to correspond to emissions of $C_i^* \approx 3000$ µgm$^{-3}$ and lower volatility compounds.“ I am not sure whether I understand this sentence correctly. Are emissions with $C_i^* \approx 3000$ µgm$^{-3}$ considered as nonvolatile? Please explain.

Page 12, line 10 and 11: This sentence should be formulated better.

Page 12, equation (27): what is $\alpha_i$?

Page 12, line 15: volatility bins are defined from 0 – 1000 µgm$^{-3}$. What is meant by the 0 µgm$^{-3}$ volatility bin? Really nonvolatile? Moreover, volatility bins do not directly refer to species. This sentence has to be reformulated.

Page 13 and Fig. 4: How is saturation concentration determined in this figure? Is it the one of the pure compound? In this case, it should be labelled with a zero.

Page 13 and Table 2: The chemical structure is not given in the manuscript. Do the species listed in Table 2 have a specific composition or only physical properties? This should be explained better.

Page 15, line28: the definition of “NMB” and “mean absolute gross error” etc. should be given in the main manuscript or in the SI.

Page 16, line 16: I state here my main criticism again: implementing LLPS just by using the LLPS threshold from You et al. (2013) does not make sense. Water associated to ions is not available for partitioning of organics. The LLPS simulation should therefore be removed.

Page 16, lines 30 – 33: “Generally, all aerosol water is expected to evaporate in an aerodynamic lens inlet used on many instruments (Zelenyuk et al., 2006), which can cause changes in the aerosol phase state (Pajunoja et al., 2016) and could potentially lead to changes in partitioning of soluble organic compounds.” Could this be the reason for the absence of the diurnal cycle in the observations? This would be an interesting point to discuss.

Page 17, Figure 10: Figure 10 should be shown also when the LLPS simulations is removed. The prediction of the occurrence of LLPS from You et al. (2013) is based on RH information and shows how relevant LLPS is in the eastern US.

Page 17, lines 4 – 7 and 20 – 23: As stated before, water uptake and semivolatile organic gas-particle partitioning including LLPS can only be modeled realistically when activity coefficients between organics and inorganic ions are included.

Page 20, lines 9 and 10: it should be explicitly stated what precursors are meant.

Page 20, line 27: what do the numbers refer to? To the different simulations or the different observations?

Page 21, lines 14 to 15: Is the second conclusion referring to the simulation $W_o > 1$? This simulation led to the largest overprediction of OA and LWC in Fig. 11! How can this be considered as the best choice?

Figure 1: this figure caption is confusing and needs improvement. The categories are not explained properly. What is “insoluble”? Can you tell how it is derived maybe by referring to an equation? The expression “other SOA” is only used in Fig. 1. Is it the same as “dry organic aerosol” from Table 2? What is “aq. SOA”? The same as “aqueous aerosol” in Table 2 or does it also include cloud water from Table 2? According to Fig. 2, there seem to be two POA species (POC and NCOM). They are considered as nonvolatile. How does this go along with the sentence: “The total nonvolatile POA in CMAQ is assumed to correspond to emissions of $C_i^* \approx 3000$ µgm$^{-3}$ and lower volatility compounds.“? This seems to imply that there is also a semivolatile POA. Is POA considered to be insoluble in water? If yes, this would be inconsistent with the water uptake associated with POA as sketched in Fig. 3.

Figure 2: the meaning of dashed and solid arrows should be stated explicitly.

Figure 3: Are the pie-charts intended to represent the contributions of the different categories to the total particulate matter realistically? Is “$P^{vap}$SOA” in Fig. 3 the same as “dry organic aerosol” in Table 2? What is meant by interaction via acidity? What is hidden within the “etc”?

Figure 4: The quality of the figure needs to be improved. The labels are on top of each other. Some of the species with large OM/OC are listed as the “dry organic aerosol” that does not seem to be involved in the uptake of water. Is this correct? If yes, why?

Technical comments:

Figures are not numbered consecutively as they appear in the text: Figure 4 is mentioned on page 5 while Fig. 3 only on page 11. Figures should therefore be reordered.

Page 4, line 15: remove first “aerosol”

Page 11; line 6: should this sentence read: “…in phase a compared to the total particulate species…”?

Page 12, line 9: “The volatility distribution of gasoline vehicle POA from May et al. (2013) and used by the CMAQ-VBS (Koo et al., 2014) was…” improve formulation.

Page 12, line 15: should it be “0.01, 0.1, 1, 10, 100, …”? Moreover, volatility bins do not directly refer to species. This sentence needs to be reformulated.

Page 14, line 6: “…not entirely…”: add “be”.

Page 15, line 20: “The standard deviation (s) in model predicted SOA fraction was much higher at 0.21 vs 0.08 in observations.” Improve formulation.

Page 16, lines 15 and 16: “Figure 9 shows how including water interactions in absorptive partitioning calculations affected model performance compared to routine monitoring networks.“ Improve formulation.

Page 17, line 14: “occurred” should be replaced by “is predicted”.

Page 17, line 19: the “ed” should be removed from “represented”.

Table 1: Is aqueous aerosol used synonymous with aqueous SOA? if yes, only one expression should be used.

Table 2: all abbreviations should be given in the table caption or a new table with all abbreviations should be added.