Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2020-396-RC1, 2020

Atmospheric Chemistry and Physics
Discussions
Open Access
EGU

Interactive comment on "Technical note: Estimating aqueous solubilities and activity coefficients of mono- and α, ω-dicarboxylic acids using COSMO-RS-DARE" *by* Noora Hyttinen et al.

Anonymous Referee #1

Received and published: 10 June 2020

This Technical Note by Hyttinen et al. introduces a new approach for the simulation of activity coefficients and aqueous solubilities of mono- and dicarboxylic acids using the COSMOtherm software. The COSMO-RS-DARE model extension is used in this work, which includes hydration and dimerization of the organic acids in water.

The paper is rather technical and focused on the choices made in this application of COSMO-RS-DARE. This is appropriate for a technical note. However, I found it difficult to judge the added value of this work in terms of the broader application of such methods in an entirely predictive way, e.g. to generate activity coefficient data sets for more complex mixtures of organic acids that have not been studied experimentally.

Printer-friendly version

Discussion paper

The manuscript does not offer much advise in this respect nor a discussion about the applications of the model. Adding a discussion about strengths and weaknesses of the introduced method would improve the paper.

The writing and structure of this technical note are overall appropriate and the topic is of interest for the ACP community. However, as indicated by several of my general and specific comments below, there are a number of mistakes, inconsistencies and unclear statements that need to be addressed before publication can be recommended.

1 General comments

- Aerosol acidity is mentioned in the introduction, but beyond that I could not find any description of the approach in COSMO-RS-DARE or COSMO-RS to account for partial dissociation of organic acids in solution. It would seem to be important as an effect that may compete with dimerization and hydration of undissociated acids. Please discuss.
- Fitting of reaction enthalpy parameters to existing experimental data was carried out, making the method perhaps less predictive than one would hope for. It is unclear how important the fitting of COSMO-RS-DARE model parameters is to achieve the presented activity coefficient and water-solubility results. If such fit parameters are essential, could you discuss the advantages of the COSMOtherm modelling approach compared to more traditional fitting of activity coefficient models for binary solutions, such as a Van Laar model or group-contribution models like UNIFAC?
- The provided quantitative comparison to available measurements of dicarboxylic acid activity coefficients or water activities is rather limited and many existing data for concentrated aqueous solutions, e.g., by Choi and Chan, J. Phys. Chem. A

2002, 106, 4566–4572 or Marsh et al. (2017), could be used for a direct comparison of measured and predicted water activities / activity coefficients of the studied binary solutions. In addition, a comparison to predictions from other models, such as the UNIFAC / AIOMFAC / E-AIM models or other such approaches would allow the reader to compare the performance of COSMO-RS-DARE to such parameterized thermodynamic models that are often used in this community.

1.1 Specific comments

- Abstract: The abstract would benefit from a quantitative statement about the average accuracy of the COSMO-RS-DARE predictions of activities or activity coefficients compared to the available experimental data and/or predictions by the standard COSMOtherm / COSMO-RS model. It would also be useful to state whether the outlined COSMO-RS-DARE method is fully predictive or not.
- Line 31: Clarify the sentence with "the Gibbs-Duhem equation was fitted"; what exactly was fitted? To my knowledge, the fundamental Gibbs-Duhem equation has no dedicated fit parameters.
- L. 32 – 34: The sentence should be improved given that, among the cited references, both Peng et al. (2001) and Choi and Chan (2002) use electrodynamic balance measurements that cover mass fractions of the solute far beyond the dilute solution range of the dicarboxylic acids. Further, the COSMOtherm predictions provided in this work (Fig. 3) seem not to be compared to such experimental data sets, even though the authors are aware of them.
- L. 35: Define abbreviations for COSMOtherm and COSMO-RS.
- L. 70: Clarification necessary; Eq. (1), (2) define activity coefficients using the pseudo-chemical potential, which differs from the "regular" chemical potential

more often used. However, comparison of Eqs. (1) and (2) raises the question how $\mu_i^{*\circ,I}$ in Eq. (1) differs from μ_i° in Eq. (2)? Eq. (2) seems to express the same relationship as Eq. (1). How exactly do they differ and what convention is used for the activity coefficients in Eq. (2)? Also, on line 71, the gas constant is expressed using kcal for energy; use of SI units would generally be preferred.

- L. 76 – 79: Clarify the basis for Eq. (3), why should that apply (references / reasons)? Also, Eq. (4) stated in the current form seems incorrect and awkward: why write the right hand side composition as $a_i^{I,\beta}(1-x_{SOL,w})$? This seems not entirely correct and to be a potential source of confusion. This confusion exists because $(1-x_{SOL,w})$ should be exactly the same as $x_{SOL,acid}$ in a binary mixture (in the same phase), yet the former expression would only be correct for binary aqueous mixtures, not in general. However, in the LLE case, the mole fractions of acid in phases α and β will differ, which is missing in Eq. (4). Why not write $a_i^{I,\beta}(x_{SOL,acid}^{\beta})$ and analogously for phase α. Further, please define the meaning of subscript SOL.
- L. 98: The following statement "by multiplying the reaction equilibrium constant with the ratio of the activity coefficients..." and need for Eq. (11) seem unwarranted and require further explanation. Why should Eq. (11) be necessary? Is this because the authors are only considering mole fractions in the reaction constant, not activities? It is unclear because based on Eq. (8) – (10), in which chemical potentials and therefore implied activity coefficients are used, there seem to be no need for Eq. (11). Are you instead using pseudo-chemical potentials? Please clarify this. Also, activity coefficients of the reactants and products in what mixture? If the mixture contains a solvent, e.g. water, which affects the values of activity coefficients, how would Eq. (11) become solvent-independent?
- L. 114: "equilibrium constants on a mole fraction basis" – Unclear what "mole fraction basis" should imply here. If the equilibrium constant is computed based

on free energies of reaction or chemical potentials of reaction, they are always dimensionless – as any thermodynamic equilibrium constant should always be (concentration-product-based equilibrium constants are only approximations and not thermodynamically correct).

- L. 130 – 133: Eq. (15) and text: It is unclear why enthalpic and entropic energy contributions should not already be accounted for by Δ^0. Also why is there a factor 2? It seems possible that the energy difference is just not known/predicted well enough, such that a fit parameter was introduced to match experimental data. Is that the motivation for the "enthalpic" contribution in Eq. (15)? Please discuss.
- L. 180 – 184: The discussion of hydration and dimerization in aqueous solutions raises the question whether the dissociation of the carboxylic acids into dissolved ions was considered in the simulation? In dilute aqueous solution, dissociation of the acid and formation of hydronium ions would seem to be favorable over non-dissociative hydration. Please discuss. Acidity and pH are mentioned in the introduction, but nothing is said about acid dissociation within COSMO-RS-DARE simulations.
- L. 198 – 199: Statement: "We used these experimental activity coefficients to fit the enthalpic parameters (c_H) for each of the acids in the COSMO-RS-DARE calculations." Given that such fit parameters were introduced, how predictive is the outlined COSMO-RS-DARE method for calculation of activity coefficients in aqueous solutions for compounds where the experimental equilibrium constants or activity coefficients are not known? Elaborating the discussion on this key point seems important for applications in atmospheric aerosol modeling. For example, if these parameters were all ignored (set to zero), how different would the predictions be for activity coefficients and solubilities in water?
- L. 214, 215: Do these calculation include phase separation for the larger acids

or is a mixed phase assumed for all compositions? At higher organic acid mole fractions, the binary mixtures show high values for water activity coefficients, as also stated in the caption to Fig. 2. This may suggest liquid-liquid phase separation could occur, which could affect the interpretation of the experimental data used for comparison. Please discuss.

- L. 261: "However, without experimental activity coefficient data of oxalic acid–water systems, we are not able to fit the c_H parameter needed for reliable estimates." – This statement is simply not correct. A literature search reveals that there exist multiple useful measurements for this binary system, such as water activities at dilute and concentrated conditions, from several sources. These include the work of Maffia and Meirelles (2001) mentioned earlier in the study, as well as work by Braban et al. (2003) and Marsh et al. (2017), which provide water activity data that should be used here.

- L. 264 and Fig. 3: The partial dissociation of dicarboxylic acids is either not considered at low acid concentrations or not discussed, even though dissociation would seem to be likely, especially for the smaller acid molecules. Was it determined to be irrelevant? This will require adequate discussion. Also, mention that mole fractions used in this work are defined assuming undissociated acid molecules (if this is indeed the case).

- L. 290 – 294: On melting temperatures and solids considered:
  "Cornils and Lappe (2000) and Omar and Ulrich (2006) also measured the melting point of oxalic acid and found temperatures almost 100 K higher than Booth et al. (2010). ..." In this context, are the differences in melting points due to unclear statements about the crystalline form of oxalic acid (anhydrous or dihydrate)? Omar and Ulrich (2006) show in their paper that the solid–solid phase transition from the dihydrate to the anhydrous oxalic acid crystal polymorph occurs around 378 K, which is close to the melting temperature stated by Booth et al., while

the anhydrous oxalic acid melts at 465 K. Therefore, the correct values of use in the COSMOtherm simulations will depend on the temperature range of interest; for room temperature, using the dihydrate from is very likely the crystalline solid to be considered and consequently the solubility equilibrium should be solved for oxalic acid dihydrate not the anhydrous form. It is also clear that the actual melting temperature and enthalpy of fusion is not as uncertain as the current text implies. Using the correct equilibrium relations for the dihydrate the agreement between COSMOtherm and the measured solubility data would be expected to be much better than shown in Fig. 3.

- L. 305: "to significantly improve the activity coefficient estimates ..." – compared to what? COSMO-RS or other methods? What about a comparison to methods like UNIFAC, such as the model by Peng et al. (2001).

- L. 308: "In addition, COSMO-RS-DARE is able to predict the miscibility of butanoic acid in water, while COSMO-RS predicts a finite solubility" – This is nice. However, how much of that success comes from fitting model parameters (c_H) rather than the use of the DARE extension?

- Table 1: State whether the listed data are for the pure anhydrous solids of the acids or for hydrates (especially in case of oxalic acid). Also replace "literature values" by more appropriate wording.

- Table 2: state the temperature for the listed equilibrium constants and for completeness also for which phase / solvent they apply (given that there is also gas phase dimerization of such acids).

- Figure 3: Why are the prediction data cut of at the solubility limits? It would seem useful to mark the solubility limit at 298.15 K for each acid, but to also show the predictions for the supersaturated range (which may apply in aerosols). This would also facilitate an extended comparison with experimental data existing for those higher concentrations, e.g. by Choi and Chan (2002) and Marsh et al. (2017). With the chosen log-scale for the x-axis, too much emphasis is put on the very dilute concentration range below $10^{-3}\ x_{acid}$, which seems not to be insightful.

2 References mentioned

Braban, C. F., Carroll, M. F., Styler, S. A., and Abbatt, J. P. D.: Phase Transitions of Malonic and Oxalic Acid Aerosols, The Journal of Physical Chemistry A, 107, 6594-6602, 10.1021/jp034483f, 2003.

Choi, M. Y., and Chan, C. K.: Continuous Measurements of the Water Activities of Aqueous Droplets of Water-Soluble Organic Compounds, The Journal of Physical Chemistry A, 106, 4566-4572, 10.1021/jp013875o, 2002.

Maffia, M. C., and Meirelles, A. J. A.: Water Activity and pH in Aqueous Polycarboxylic Acid Systems, Journal of Chemical & Engineering Data, 46, 582-587, 10.1021/je0002890, 2001.

Marsh, A., Miles, R. E. H., Rovelli, G., Cowling, A. G., Nandy, L., Dutcher, C. S., and Reid, J. P.: Influence of organic compound functionality on aerosol hygroscopicity: dicarboxylic acids, alkyl-substituents, sugars and amino acids, Atmos. Chem. Phys., 17, 5583-5599, 10.5194/acp-17-5583-2017, 2017.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-396, 2020.