

1 We thank both reviewers for their constructive comments which we think have helped us improve the
2 manuscript. Our point-by-point responses to both reviewer comments along with the corresponding
3 changes we have made to the revised manuscript (attached in the end of the letter with the changes
4 highlighted) are given below. Besides the revisions requested by the reviewers we have modified Sect.
5 2.2.4 according to the corrigendum we published as an author comment, as well as corrected some typos
6 and small inclarities we found in the original manuscript.
7

8 **Response to Reviewer #1**

9

10 We thank this Reviewer for the thoughtful read of our manuscript and the insightful comments. Our
11 point-by-point responses to the Reviewer's comments are below. The direct quotes from the Reviewer
12 are in *italics* and our responses in normal font.
13

14 *(1) I am very unsure of, and uncomfortable with, the use of the terminology Basis Set in the context of*
15 *the current study. It makes no sense in terms of the common usage in theoretical and computational*
16 *chemistry, nor that in linear algebra. Its usage in terms of the well-established VBS was also difficult to*
17 *understand and its adoption in the current work by analogy to the VBS requires explanation. As stated,*
18 *the current manuscript considers "a continuous distribution of solubilities" in classes one decade apart.*
19 *This provides a suitable reference scale on an appropriate axis, but I do not see where there is any*
20 *reference "set" (of functions, of compounds, of properties, of vectors etc...) that can be used to make*
21 *sense of the impacts of particle component solubility. This is particularly important in the context of the*
22 *non-independence of the presented component description in terms of solubility and the extended*
23 *volatility description where the second dimensions considers the degree of oxygenation (in terms of O:C*
24 *or oxidation state), see below.*
25

26 We understand the point of the reviewer and indeed the different dimensions of the "basis sets" are not
27 expected to be independent of each other. We will therefore primarily use the term "solubility
28 distribution" instead of "solubility basis set (SBS)" throughout the revised manuscript, except for the
29 point in which the analogy to the widely-used VBS approach is discussed.
30

31 *(2) To expand on this point, given the relationship between aqueous solubility of organic molecules and*
32 *their polarity (and hence practically, their oxygen-containing functionality), the solubility consideration*
33 *in the current paper is clearly not independent of the 2-D VBS of Donahue et al., 2011 or the Carbon*
34 *number - oxidation state representation of Kroll et al., 2011. Whilst both approaches are mentioned in*
35 *the current work, the relationship between the approaches is insufficiently discussed. It appears that the*
36 *current work concerns itself with the cloud droplet activation behaviour of OA components, whereas*
37 *the "VBS" approaches are more concerned with formation and transformation of the OA. However, the*
38 *relationship between degree of oxygenation and hygroscopicity has been widely investigated, both in*
39 *terms of the VBS and otherwise (e.g. in terms of AMS m/z 44 in the paper introducing the 2D-VBS).*
40 *Introduction of yet another approach without contextualisation appears to add to the confusion, rather*
41 *than to its clarification.*
42

43 It is expected that the mixture solubility distribution will depend on both polarity and thus O:C ratio
44 (Donahue et al., 2011; Kroll et al., 2011) as well as the molar mass (Shiraiwa et al., 2014) of the mixture
45 constituents. To illustrate this, we have added a figure depicting an example solubility distribution for a
46 representative case of α -pinene SOA (Chen et al., 2011), where the solubilities have been estimated
47 using the SPARC (see e.g. Wania et al., 2014 and references therein) prediction tool. Also the mean
48 molecular mass and O:C ratio for each solubility bin are presented. A brief discussion of the expected
49 dependence of the solubility on these different molecular properties has been added along with the
50 description of the new figure to the revised manuscript.
51

52 *(3) Line 7 p28530 it is stated that "This assumption is justified to a first order by the different*
53 *equilibration time scales of the droplets with respect to water vapour and the organic vapours in typical*

1 atmospheric conditions" when referring to the lack of consideration of partitioning of organics between
2 the gaseous and either aqueous or non-aqueous phases in the current study. This is a bold statement.
3 The activation is driven by the increase in water saturation ratio by cooling (updraught in a cloud,
4 radiative or advective in a fog). The (low) rate of change of saturation ratio caused by this will lead to
5 the mass flux of water. Since the number of water molecules is high, it is frequently assumed that the
6 number of collisions leads much more rapid condensation of water vapour than of other components.
7 However, the rate of change of saturation ratio of organic molecules can be very much higher (owing
8 to both temperature, but equally or more importantly photochemical reaction). Since the vapour
9 pressure of the organic can be many orders of magnitude lower than water, a strong diffusion gradient
10 between the gas phase and the particle surface can be very rapidly established. There are very many
11 potential oxygenated organic compounds that can all be rapidly produced in response to strong
12 emission and photochemical changes. It is far from clear to me that the addition of soluble mass to a
13 particle in moving towards cloud base can be ignored to a first order approximation (and hence that
14 the solubility distribution of components at lower RH will be the same as the solubility distribution of
15 components close to activation). The Topping et al. Nature Geoscience paper (doi:10.1038/ngeo1809)
16 should be referenced in this regard. Again, the blurred relationship between volatility treated in the
17 "VBS" approaches and solubility treated here is related to this. It appears conceptually peculiar to silo
18 treatments of partitioning and water uptake into the separate approaches; the moist atmosphere is a
19 single entity in which both processes simultaneously occur.

20
21 We agree that this is an important assumption and that it deserves additional discussion. It is also true
22 that for a comprehensive picture of the CCN activation, all the relevant dynamic processes for all the
23 involved species (and their gas phase concentrations) need to be explicitly considered. This requires,
24 however, a much more complex modeling framework and detailed information about the connections
25 between volatility and water solubility of the different organic compounds that is not currently available.
26 For the scope of this work we felt that it is important to focus on the solubility effects on the CCN
27 activation process, without adding the uncertainty related to e.g. the gas-phase concentrations of the
28 organic species and atmospheric dynamics. This is clearly a first step and investigations with a cloud
29 parcel model including all the relevant dynamic processes will be carried out in a future study. We think
30 that this approach is justified, as this is the first study investigating the sensitivity of the CCN activation
31 processes to a wide selection of solubility distributions. In the revised manuscript we have added a
32 reference to Topping et al. (2012) and replaced the statement of the justification of this assumption with
33 a more nuanced discussion of the corresponding issues mentioned by the reviewer.

34
35 **(4)** *I'm a little confused in how the approach described here to calculate the equilibration of*
36 *components between the insoluble core and the surrounding aqueous shell relates to previous*
37 *approaches to deal with phase separation in complex mixtures in atmospheric aerosol (notably that*
38 *presented by Topping et al., 2013, doi:10.1039/C3FD00047H, which should be referenced). The*
39 *solubility in the non-aqueous and aqueous phases must implicitly account for non-ideality (and e.g.*
40 *degree of dissociation of weak acids and bases) and hence must be considered an effective solubility in*
41 *the mixture. In the aqueous phase, this effective solubility may or may not be relatively simply related*
42 *to the infinite dilution solubility of the component. In any case, it will very likely be relative humidity*
43 *dependent. As such, the effective solubility distribution in the two phases will be RH dependent and this*
44 *will (potentially significantly) affect the shape of the Kohler curves and the results illustrated in Figure*
45 *3. Coupled to the equilibration between the gas phase and the two condensed phases, it might be*
46 *expected that this effect is substantial. In figure 8 a simple sensitivity to an assumed activity relationship*
47 *is explored (using equation 5). Such calculations can be carried out for simple real mixtures where the*
48 *activity coefficient data (albeit on a mole fraction scale) are available (and complex mixtures where the*
49 *activity coefficients can be predicted). This sensitivity is assumed to constrain the real behaviour (very*
50 *likely correct). It would be useful to know whether the real behaviour more normally follows one or the*
51 *other of these limiting cases. Even in the absence of confirmatory experimental data, such calculations*
52 *are possible.*

53

1 We have added the reference to Topping et al. (2013) and the corresponding discussion. We have also
2 cited the work of Cappa et al. (2008) with some discussion on the potential non-ideality of
3 multicomponent mixtures of dicarboxylic acids. Unfortunately, there is very little experimental
4 information available about atmospheric organic compound activity coefficients in the aqueous or
5 multicomponent organic phases. Data is available for individual organic compounds, but it is not clear
6 how useful these data would be to estimate the mixture behavior. Our simplified ideality assumption for
7 the aqueous phase activities below the saturation concentrations is a reasonable starting point. The
8 corresponding solubilities discussed here should be interpreted, as the reviewer suggests, as “effective”
9 solubilities. A brief discussion has been added to address this point in the revised manuscript.

10
11 *(5) It is stated on line 15 of 28532 that "... all the equations presented below can be rederived in a*
12 *relatively straightforward manner taking into account a potential difference between the mole and the*
13 *mass fractions in the organic phase". This is correct, but the reference scale difference for the activity*
14 *coefficients (mass or mole fraction) must be emphasised. This must be made explicit - it is the same*
15 *problem as encountered in the VBS and is more important here, since the molecular weight distribution*
16 *and degree of functionalisation will significantly affect effective solubility and hence phase partitioning.*

17
18 This is correct. We have added a sentence to the revised manuscript emphasizing the fact that we have
19 studied the special case where the mole and mass fraction scales are the same, but that these should be
20 distinguished for realistic atmospheric mixtures.

21
22 **Minor:**

23
24 *(6) Some reference to the expected individual component effective solubility range for atmospheric*
25 *components would be helpful. The very heavy functionalisation (of multifunctional peroxides,*
26 *peroxyacids etc...) expected and required to attain observed O:C ratios with a measured molecular*
27 *weight distribution of organics provides some surprisingly high solubilities. Along with the very low*
28 *solubilities of slightly functionalised high molecular weight primary compounds, this will lead to a very*
29 *wide distribution of solubilities as shown in Figure 2. Are there any data available to say whether the*
30 *real atmosphere is likely to be closer to one or another of the chosen distributions? If not, what are the*
31 *difficulties and are there suggested routes forward?*

32
33 To link our approach better to realistic atmospheric organic mixtures, we have now constructed the
34 solubility distribution for an example composition of secondary organic aerosol (SOA) mixture derived
35 from α -pinene oxidation (see Chen et al., 2011 for details of the gas-phase composition and conditions
36 corresponding to this distribution). It should be noted though that this is only one example, and the
37 organic aerosol formed under different conditions or derived from other precursors can have vastly
38 different solubilities, nevertheless being equally well represented by the solubility distribution approach.
39 The results from our manuscript can then further be used to assess the requirements (e.g., range of
40 solubilities, resolution) of a solubility distribution measurement approach to represent the CCN
41 activation of these mixtures. We have added discussion on the link of the study to realistic atmospheric
42 organic mixtures, along with adding a new figure depicting an estimated realistic solubility distribution
43 to the revised manuscript.

44
45 *(7) I am not sure that the 2 sentences in the abstract from line 17 to 21 and 21 to 24 are not contradictory.*
46 *Please check. Can the stated material in the first sentence be assumed completely soluble and is this*
47 *material not just a subset of the material considered within the distribution mentioned in the second*
48 *sentence and hence requiring understanding of the details of its distribution in solubility? In any case,*
49 *the arguments considered together in such close proximity before the reader has accessed the rest of*
50 *the paper are too complicated for easy interpretation.*

51
52 Indeed these two statements can cause some confusion. We have revised the statements to read:

1 ”Depending on what was assumed about the organic phase activity, material with solubilities larger than
2 0.1-100 g L⁻¹ could be treated as completely soluble in the CCN activation process over atmospherically
3 relevant particle diameters and supersaturations. Our results indicate that understanding the details of
4 the solubility distribution is thus probably necessary only in the range of 0.1 to 100 g L⁻¹, while
5 resolution outside this solubility range will not add much knowledge to understanding the CCN
6 activation of the mixture.”
7

8 **(8)** *On p28526, 3 prior assumptions are described that are considered in the paper. The kappa*
9 *consideration is not the only single parameter representation of hygroscopicity (e.g. Wex et al., 2007,*
10 *doi:10.1029/2006GL028260), nor the first (e.g. Rissler et al., ACP, 2004). These treatments should be*
11 *acknowledged.*

12
13 Thank you for pointing this out, we will correct this in the revised manuscript.
14

15 **(9)** *It isn't clear how the presence of inorganics effects results in this paper. It is likely that there will be*
16 *an inorganic fraction under most expected atmospheric conditions and this should be mentioned in the*
17 *discussion.*

18
19 This is an important point as the present work addresses explicitly only organic particles. The work can
20 be extended to include inorganic salt adding one more variable for consideration (the organic mass
21 fraction in the particle). However, accurate considerations on the mixture activity and potential phase-
22 separation effects will become much more complex, requiring detailed experimental data as well as
23 advanced thermodynamic models – being thus a very interesting topic for a future study. We will add a
24 brief discussion on this to the revised version of the manuscript.
25

26 **References**

27
28 Cappa, C. D., Lovejoy, E. R., and Ravishankara, A. R.: Evidence for liquid-like and nonideal behavior
29 of a mixture of organic aerosol components, *Proc. Natl. Acad. Sci. U. S. A.* 105, 18687, 2008.
30

31 Chen, Q., Liu, Y., Donahue, N. M., Shilling, J. E., and Martin, S. T.: Particle-phase chemistry of
32 secondary organic material: Modeled compared to measured O:C and H:C elemental ratios provide
33 constraints,” *Environ. Sci. Technol.*, 45, 4763, doi:10.1021/es104398s, 2011.
34

35 Donahue, N. M., Epstein, S. A., Pandis, S. N., and Robinson, A. L.: A two-dimensional volatility basis
36 set: 1. organic-aerosol mixing thermodynamics, *Atmos. Chem. Phys.*, 11, 3303-3318, doi:10.5194/acp-
37 11-3303-2011, 2011.
38

39 Kroll, J. H., et al.: Carbon oxidation state as a metric for describing the chemistry of atmospheric organic
40 aerosol, *Nature Chem.* 3, 133, 2011.
41

42 Shiraiwa, M., Berkemeier, T., Schilling-Fahnestock, K. A., Seinfeld, J. H., and Pöschl, U.: Molecular
43 corridors and kinetic regimes in the multiphase chemical evolution of secondary organic aerosol, *Atmos.*
44 *Chem. Phys.*, 14, 8323-8341, doi:10.5194/acp-14-8323-2014, 2014.
45

46 Wania, F., Lei, Y. D., Wang, C., Abbatt, J. P. D., and Goss, K.-U.: Novel methods for predicting gas-
47 particle partitioning during the formation of secondary organic aerosol, *Atmos. Chem. Phys.* 14, 13189,
48 2014.
49

50 Topping, D. O., Connolly, P., and McFiggans, G.: Cloud droplet number enhanced by co-condensation
51 of organic vapours, *Nature Geosci.*, 6, 443, 2012.
52

1 Topping, D.O., Barley, M., and McFiggans, G.: Including phase separation in a unified model to
2 calculate partitioning of vapours to mixed inorganic–organic aerosol particles, *Faraday Discuss.*, 165,
3 273, 2013

4 5 **Response to Reviewer #2**

6
7 We thank the reviewer for the constructive comments. Our point-by-point responses to the Reviewer's
8 comments are below. The comments of the Reviewer are in *italics* and our responses in normal font.

9
10 *(1) The idea of the solubility basis set is novel and promising. It excites interest in the possibility to*
11 *develop accurate computationally efficient parameterizations of the evolution of solubility for large*
12 *scale models. The study at hand presents only limited calculations that probe parts of the parameter*
13 *space.*

14
15 The proposed treatment of the organic aerosol solubility distribution does simplify the multiphase
16 thermodynamics. However, we have covered most of the atmospherically relevant parameter space
17 related to the coupling of organic mixture solubility and CCN activation, namely:

- 18 1) the characteristics of atmospherically relevant solubility distributions (number of compounds,
19 solubility range, shape of distribution and assumption about organic phase activity);
- 20 2) the sizes of the particles that can act as CCN;
- 21 3) atmospheric supersaturations.

22
23
24 While it is true that we did not cover the range of other thermodynamic properties (e.g., densities, molar
25 masses as well surface tensions) of the organics, it was a conscious choice to isolate the solubility-related
26 effects, which we chose to be the scope of this study – in addition to the introduction of the solubility
27 distribution framework. Including the variation of all the other relevant thermodynamic properties of
28 the organic compounds would be an interesting topic for future study. Furthermore, as also pointed out
29 by Reviewer #1, to fully assess the impact of solubility on cloud formation, the results presented here
30 need to be coupled to a model accounting for gas-particle and surface-bulk partitioning of the organics
31 as well as atmospheric dynamics and the condensation kinetics of water vapor – framing out another
32 direction for future work. We have clarified the justification for the chosen scope and added discussion
33 about the potential future directions along these lines in the revised manuscript.

34
35 *(2) Although the proposed treatment seems to capture the important physics, there are errors in the*
36 *theory that require correction.*

37
38 The major issue is related to the use of the solution density in the Kelvin term (see below). We have
39 now repeated the calculations according to the suggestions of the reviewer and found that its impact on
40 the presented results is small and does not affect the conclusions of the study. The manuscript has been
41 revised accordingly.

42
43 *(3) The presented main finding that is advertised - knowledge of solubility in the range of 0.1 and 100*
44 *g L⁻¹ is most critical for modeling CCN - is already well known from past studies (e.g. Hori et al., 2003,*
45 *J. Aerosol Sci.) and insufficient to merit publication of a new paper on the subject.*

46
47 The Hori et al. (2003) study and a number of similar previous studies have focused on a small number
48 of specific compounds – thus investigating only specific points in the parameter space. To our
49 knowledge, previous studies have not systematically explored such a comprehensive set of possible
50 solubility distributions and atmospheric conditions, including the consequences of mixture effects on
51 solubility. This is of course something that is only achievable through theoretically generated mixtures
52 like ours, but is also among the main strengths of our approach. Additional important contributions of
53 our work include:

- 1 1) The introduction of the solubility distribution as a framework for representing complex organic
2 aerosol solubility.
- 3 2) The finding that material with effective solubility above $0.1\text{-}100\text{ g L}^{-1}$ behaved as completely
4 soluble for most of the approximately 6000 unique combinations of solubility distributions
5 (representing 144 unique mixtures), particle sizes and supersaturations studied. The most
6 important piece of information to know in this regard is the thermodynamics of the insoluble
7 organic phase.
- 8 3) For the large range of organic mixture solubilities, particle dry diameters and supersaturations
9 explored, the single-parameter approaches (using single ε or κ value to describe the solubility of
10 the mixture) were generally sufficient to predict the activation diameters of the mixture aerosols
11 within 10%.
- 12 4) Our results also contribute to unraveling the reasons behind the "apparent simplicity" of
13 atmospherically-relevant complex organic mixtures (e.g. Engelhart et al., 2011). Based on our
14 results it is clear that even with vastly different solubility distributions one can yield very similar
15 CCN-activation behavior (and consequently values of κ or ε), as the parameter that matters is the
16 material above c_t . This result is perhaps not surprising but it has not been systematically probed for
17 such a large number of unique mixtures.

18
19 We have modified the abstract and conclusions of the revised manuscript to better highlight these
20 findings and their novelty in relation to previous work, adding also reference to the work by Hori et al.,
21 2003.

22
23 *(4) The evaluation of the full model against the ideal mixture, kappa-model and epsilon-model seem not*
24 *very relevant to me, partially attacking and deconstructing a strawman.*

25
26 We respectfully disagree, although we understand that we are essentially comparing what is supposed
27 to be a simplification of the more complex theory to the original (more complex) description. First of
28 all, we quantitatively assess how large error in the activation diameter one makes in assuming these
29 simplified solubility descriptions instead of the "full" picture, putting the importance of solubility into
30 the context of the other uncertainties related to cloud-formation process. The results convincingly show
31 that in many cases, only one parameter (either κ or ε) is definitely enough to accurately calculate the
32 activation diameter in e.g. climate models. The error introduced by this simplification is negligible
33 compared to other uncertainties. Second, while the results make physical sense, we do not think that
34 they can be deduced without a systematic study like ours. For instance, the importance of the assumption
35 on the organic phase interactions – which turns out to be quite important in defining c_t (and thus highly
36 relevant for linking e.g. experimentally determined ε or κ to real mixture properties) at given conditions
37 – is something that has not been clearly pointed out before.

38
39 *(5) Instead, the most pressing question, how to effectively model the 50% point of partitioning (c_t) in*
40 *relationship to the aqueous phase formed by a mix of species with different hygroscopicities (including*
41 *inorganic compounds and the dependence of the number of components) is not at all or insufficiently*
42 *explored. Not varying the molecular weight and not including inorganic compounds seems to miss the*
43 *most important question: how does the aqueous phase (number of substances and their hygroscopicity)*
44 *influence c_t for a given solubility basis set? If inorganic or higher hygroscopic organics are present in*
45 *the mix, c_t should shift to lower values. In my opinion, quantifying shifts in c_t in the context of the SBS*
46 *and hygroscopicity of the aqueous phase ought to be the main focus of this work.*

47
48 Indeed, as also indicated by the distributions present in Fig. 9 in the present paper, the c_t depends also
49 on the number of components, supersaturation, particle size and importantly what is assumed about the
50 organic phase thermodynamics. We have added a brief discussion of these points to the revised
51 manuscript as well as a new figure illustrating the dependence of c_t on the number of components present
52 in the mixture as well as the molar masses of these species. These new results show that what by far
53 matters the most is the assumption about the organic phase activity. In the case of the unity activity

1 assumption, the dependence on the number of components is in practice the same behavior as observed
2 in the distributions presented in Fig. 8 of the original manuscript.

3
4 While it would be interesting to do a specific study on the impact of inorganic compounds on c_i , we feel
5 that it is out of the scope of this paper, whose target is to lay out the general framework and behavior of
6 a large number of different solubility distributions rather than look at any specific mixtures. To the
7 extent that we think it fits the scope of this work, we believe that the solubility distributions that have
8 been probed also represent well the case where an inorganic component would be present – particularly
9 as combined with demonstrating the sensitivity of the results to the molar mass. Furthermore, we have
10 added discussion on the dependence on the c_i on molar mass and O:C ratio through including an example
11 of atmospherically relevant organic aerosol mixture accompanied by a new figure.

12
13 *(6) I suggest to include a discussion paragraph on the vision how the solubility basis may be useful
14 beyond the immediate sensitivities that are being explored in this manuscript. It may be useful to explore
15 this in the context of experimental data showing that the OA hygroscopicity increases are driven by
16 evolving solubility with chemical aging as discussed in Kuwata et al., 2012, ACP and Suda et al., 2014
17 ES&T. This discussion should include computational considerations. For example, what is the current
18 (or expected gain) from running a 10, 100, 1000 component solubility model vs a model that has could
19 predict c_i , and κ for the aqueous phase. I performed a quick test running a solubility model with
20 unique solubility, molecular weight, and mixing fraction and CPU time only increased when reaching
21 10000 unique components. Presumably the purpose the SBS to be (much) more computationally efficient
22 and should comment and factors influencing computational speed.*

23
24 This is a good idea and we have added the corresponding brief discussion about the future use of the
25 proposed framework to the revised manuscript. We also believe that the solubility distribution can be a
26 useful approach for modeling how the mixture hygroscopicity evolves with atmospheric age, linking it
27 to the chemical properties of the molecules present in the mixture. In this regard, we have added
28 references to the work by Kuwata et al. (2012) and Suda et al. (2014). Furthermore, as also stated above,
29 we believe that the solubility distribution framework can yield useful insights in the interpretation of
30 laboratory data for complex organic mixtures – in particular on the apparent simplicity of their CCN
31 activation despite the well-known molecular complexity of the mixtures. Furthermore, the solubility
32 distribution approach will probably be useful in studies investigating the relative importance of
33 solubility versus semi-volatile partitioning (e.g. Topping et al., 2012) and adsorption effects (e.g. Kumar
34 et al., 2009) for water uptake and CCN activation.

35
36 We have also added a brief discussion on the computational cost to the revised manuscript. Indeed,
37 10000 compounds is the limit where significant increase in the required computational time is expected
38 for the simple Köhler-theory -based CCN-activation calculation. However, for large-scale models, every
39 additional simulated species (in this case solubility bin) can increase the computational cost significantly
40 due to the need of simulating its atmospheric chemistry, transport, removal, etc. The resolution needed
41 for accurate treatment will require better understanding of the actual solubility distributions and their
42 linkages with other parameters (e.g., volatility, O:C, etc.).

43
44 *(7) Although there is nothing really wrong with it per se, the adoption of mass units deviates from what
45 has been done in the past and seems awkward to me. It doesn't make the theory novel and may lead to
46 unnecessary confusion when comparing against other manuscripts. The new part is the solubility basis
47 set on which the focus should lie.*

48
49 We assume that the Reviewer is referring to the use of mass units in expressing the solubilities – as for
50 the mixture composition we have chosen a case where the mass and the molar units are in fact equivalent.
51 To our knowledge expressing pure-component solubilities in g L^{-1} (or in mass per mass of water) is
52 standard in e.g. chemical engineering handbooks, so it is difficult to know what (and which past work)
53 the Reviewer is referring to. We would therefore prefer to keep the current units and also add some text
54 clearly acknowledging the difference between molar and mass-based units.

1
2 **(8)** *Eq. (1): The use of solution density in the Kelvin equation is incorrect. See Kreidenweis et al.*
3 *(2005, ACP, pg. 1359). Discussion regarding the use of solution density should be removed.*
4

5 This is true as we are assuming an ideal aqueous phase. We have now corrected this in all the calculations
6 and revised the manuscript accordingly.
7

8 **(9)** *Surface tension (Table 3, Eq. 1, pg. 28536): “calculating the surface tension as weight averages of*
9 *the water and pure organic values”. I am unaware for the basis for that treatment. The surface tension*
10 *of the aqueous solution usually follows a logarithmic law (e.g. the Szyszkowski equation) derived from*
11 *Gibbs adsorption isotherms. Furthermore, application of surface tension values derived from bulk*
12 *solutions produces incorrect results since the surfactant cannot lower the free energy of the surface and*
13 *contribute to the water activity simultaneously. A correct theoretical treatment will have to include*
14 *accurate partitioning between the phases (Sorjamaa et al., ACP, 2004). Since surfactant treatment is*
15 *not at the center of this work, I recommend removing it completely and assuming the surface tension of*
16 *pure water.*
17

18 This is a valid concern. The linear weighing for mixture surface tensions is typically applicable only for
19 hydrocarbon mixtures or other mixtures with relatively similar components (see e.g. Poling et al., 2001),
20 and indeed more sophisticated approaches for aqueous solutions are based on correction terms applied
21 to the surface tension of pure water (see for example Werner et al. 2014 and references therein). We
22 have chosen the linear weighing in an effort to introduce a simple approach for the lowering effect of
23 the organics on the water surface tension – particularly taking into account the fact that we used only
24 one representative value for the organic surface tension. On another note, the linear approach turns out
25 to be a reasonable approximation for succinic acid, although it underestimates the impact of the acid on
26 the mixture surface tension (see Werner et al. 2014 and references therein). However, to avoid this kind
27 of complexity, we have redone all the calculations assuming the surface tension of water and have
28 updated the figures in the revised manuscript accordingly.
29

30 **(10)** *Eq. (5): the units are a mix between mass and molar (csat is g/g and Yi is mole fraction?*
31 *Maybe it is just a typo in the text below?).*
32

33 Equation 5 is correct, since analogously to the gas-liquid equilibrium the organic-aqueous phase
34 equilibrium is dependent on the mole fraction in the insoluble organic phase, as long as the activity
35 coefficient is also defined on a molar basis. In principle, similar formulation could be made with mass-
36 based definitions. In our case, however, the two approaches are equivalent due to the assumption of
37 constant molar mass throughout the organic mixture. We have clarified this in the revised manuscript.
38

39 **(11)** *The statement that “water -solubility determines the aerosol water content at equilibrium” is wrong*
40 *or only partially correct, depending on how one views the problem. The water content at equilibrium is*
41 *controlled by RH, surface tension, and the molecular weight and activity of the dissolved components.*
42 *If all components are dissolved, solubility exerts zero control on aerosol water content (as*
43 *acknowledged in the paper).*
44

45 True. We have rephrased this sentence in the revised manuscript.
46

47 **(12)** *The distinction between ideal mixture vs unit activity model is unclear. In an ideal solution, the*
48 *activity coefficient of each component is equal to unity. Thus, the two models are semantically identical.*
49 *Despite several attempts I am unable to extract from the paper how the two cases are different.*
50

51 In the ideal mixture the activity coefficient is unity, meaning that the effective solubility of a given
52 component i scales with the composition of the insoluble organic as indicated by Eqs. 1 and 2. In the
53 unity activity case, on the other hand, the mixture activity (the product of activity coefficient and mole

1 fraction) is unity – indicating that the individual components dissolve to water in the same way as they
2 would if they were present in their pure form. We have added this clarification to the revised manuscript.
3

4 **(13)** *Comparing the limited solubility model to the full dissolution model is a straw man argument. Of*
5 *course they will agree if all species will dissolve, disagree if they do not, and the disagreement will be*
6 *proportional to c_t . It is trivial that if one breaks the assumption of a model that it won't work. I suggest*
7 *removing that discussion from the paper.*
8

9 We understand the point that the Reviewer is making, but think there is indeed value in quantifying the
10 actual differences in activity diameter predictions – given that we have explored a large fraction of all
11 atmospherically relevant cases. This helps in putting the importance of solubility in context with other
12 potential uncertainties present in modeling the atmospheric cloud formation processes. What we also
13 think is remarkable in our results is that assuming a single c_t for all the mixtures yields a reasonably
14 good prediction of the soluble fraction and thus the activation diameters.
15

16 **(14)** *As far as I can tell, the epsilon model and the kappa model are identical since $\kappa =$*
17 *$\epsilon \cdot \kappa_{max}$ and κ_{max} is fixed in the study. Therefore figures 7a,b are redundant, as are*
18 *figures 11, c,e. There may be merit to keeping both treatments if molecular weight of the organic is*
19 *allowed to vary.*
20

21 This is partly true, as κ or ϵ are directly related. However, the sensitivity of the activation diameter to
22 these parameters is different, since when the ϵ model is used, more information needs to be given in the
23 fitting process. However, the results shown in Fig. 6 of the discussion paper show that this information
24 does not make a big difference in the predicted activity diameters – although there is a small effect seen
25 in the narrower quartile range present for the ϵ model.
26

27 **(15)** *Figure 7b suggests that there are numerical issues with the model. How can the fitted kappa exceed*
28 *kappa max?*
29

30 This was due to using the solution density and surface tension in the Kelvin term of the full model
31 calculations while this information was not given to the simplified schemes. As suggested by the
32 reviewer, we have now repeated all the calculations assuming the surface tension and density of pure
33 water, and have modified the related figures accordingly.
34

35 **(16)** *It would be helpful to include isolines of epsilon in Figure 3,b to visualize the range of solubilities*
36 *probed by the model.*
37

38 This is a good suggestion and we have redrawn the corresponding figure adding the suggested isolines.
39

40 **(17)** *Page 28534, line 5 to end of paragraph: Is it necessary to list how many Köhler curves were made?*
41 *More interesting would be e.g. calculations of $d(\text{property})/d(\text{other property})$ in your calculations.*
42

43 Indeed, the total number of the Köhler curves is perhaps not that interesting, but reason we have included
44 the number of the activation points just to illustrate the number of different combinations of solubility
45 distributions, particle sizes and supersaturations we investigated. We are not sure what the reviewer
46 means by calculations of $d(\text{property})/d(\text{other property})$, and would prefer to keep the description as it is.
47

48 **(18)** *Page 28525, Typo line 11, "Aalbrecht"*
49

50 We have corrected the typo.
51

52 **(19)** *Page 28525, Line 19-25: connect with "and" ?*
53

54 We have corrected the typo.

1
2 (20) Page 28526, Line 5, “than pure compounds”
3

4 We have modified the revised manuscript accordingly.
5

6 (21) Page 28526, Line 27, *kappa* does not strictly vary between 0 and 1, I think this is
7 misleading as it suggests that *kappa* is a ratio.
8

9 We have clarified this in the revised manuscript.
10

11 (22) Page 28526, Paragraph beginning on line 20: I think this paragraph oversimplifies the different
12 representations of solubility in CCN activation studies. Furthermore Table 1 does not aid in the
13 understanding of the paragraph or the simplifying assumptions. I would recommend removing Table 1
14 and expanding the paragraph.
15

16 We have expanded the paragraph as suggested. However, without further information it is difficult to
17 know what oversimplification the Reviewer refers to, and would prefer to keep the table as it is.
18

19 (23) Section 2.1.1: The section heading does not reflect the content.
20

21 We have revised the title to read “Equilibrium between water vapour and an aqueous phase containing
22 dissolved material”
23

24 (24) Page 28540, Line 17, change “practise” to “practice”
25

26 Corrected.
27

28 (25) Pg 28543, Line 4: “correlaction” to “correlation”
29

30 Corrected.
31

32 **References**

33
34 Engelhart, G. J., Moore, R. H., Nenes, A., and Pandis, S. N.: Cloud condensation nuclei activity of
35 isoprene secondary organic aerosol, *J. Geophys. Res.*, 116, D02207, doi:10.1029/2010JD014706, 2011.
36

37 Hori, M., Ohta, S., Murao, N., and Yamagata, S.: Activation capability of water soluble organic
38 substances as CCN, *Aerosol Sci.*, 34, 219, 2003.
39

40 Kumar, P., Sokolik, I. N. and Nenes, A.: Parameterization of cloud droplet formation for global and
41 regional models: including adsorption activation from insoluble CCN, *Atmos. Chem. Phys.*, 9, 2517,
42 doi:10.5194/acp-9-2517-2009, 2009.
43

44 Poling, B. E., Prausnitz, J. M., and O’Connell, J. P.: *Properties of gases and liquids*, McGraw-Hill,
45 2001.
46

47 Topping, D. O., Connolly, P, and McFiggans, G.: Cloud droplet number enhanced by co-condensation
48 of organic vapours, *Nature Geosci.*, 6, 443, 2012.
49

50 Werner, J., Julin, J., Dalirian, M., Prisle, N., Öhrwall, G., Persson, I., Björneholm, O., and Riipinen, I.:
51 Succinic acid in aqueous solution: connecting microscopic surface composition and macroscopic surface
52 tension, *Phys. Chem. Chem. Phys.* 16, 21486, 2014.
53
54

1 **Connecting the solubility and CCN activation of complex**
2 **organic aerosols: A theoretical study using solubility**
3 **distributions**

4
5 I. Riipinen^{1,2}, N. Rastak¹ and S. N. Pandis^{2,3}

6 ¹ Department of Environmental Science and Analytical Chemistry, Stockholm University,
7 Stockholm, Sweden

8 ² Center of Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA, United
9 States

10 ³ Department of Chemical Engineering, University of Patras, Patras, Greece

11 Correspondence to: I. Riipinen (ilona.riipinen@aces.su.se)

12

13 Abstract

14 We present a theoretical study investigating the cloud activation of multicomponent organic
15 particles. We modeled these complex mixtures using solubility distributions (analogous to
16 volatility distributions in the VBS, i.e. volatility basis set, approach), describing the mixture as
17 a set of surrogate compounds with varying water-solubilities in a given range. We conducted
18 Köhler theory calculations for 144 different mixtures with varying solubility range, number of
19 components, assumption about the organic mixture thermodynamics and the shape of the
20 solubility distribution, yielding approximately 6000 unique cloud condensation nucleus (CCN)-
21 activation points. The results from these comprehensive calculations were compared to three
22 simplifying assumptions about organic aerosol solubility: 1) complete dissolution at the point
23 of activation; 2) combining the aerosol solubility with the molar mass and density into a single
24 effective hygroscopicity parameter κ ; 3) assuming a fixed water-soluble fraction ε_{eff} . The
25 complete dissolution was able to reproduce the activation points with a reasonable accuracy
26 only when the majority (70-80%) of the material was dissolved at the point of activation. The
27 single parameter representations of complex mixture solubility were confirmed to be powerful
28 semi-empirical tools for representing the CCN activation of organic aerosol, predicting the
29 activation diameter within 10% in most of the studied supersaturations. Depending mostly on

1 the condensed-phase interactions between the organic molecules, material with solubilities
2 larger than about 0.1-100 g L⁻¹ could be treated as soluble in the CCN activation process over
3 atmospherically relevant particle dry diameters and supersaturations. Our results indicate that
4 understanding the details of the solubility distribution in the range of 0.1 to 100 g L⁻¹ is thus
5 critical for capturing the CCN activation, while resolution outside this solubility range will
6 probably not add much information except in some special cases. The connections of these
7 results to the previous observations of the CCN activation and the molecular properties of
8 complex organic mixture aerosols are discussed. The presented results help unravel the
9 mechanistic reasons behind observations of hygroscopic growth and CCN activation of
10 atmospheric secondary organic aerosol (SOA) particles. The proposed solubility distribution
11 framework is a promising tool for modeling the interlinkages between atmospheric aging,
12 volatility and water-uptake of atmospheric organic aerosol.

13

14 **1 Introduction**

15 Interactions of atmospheric aerosol particles with ambient water vapour determine to a large
16 extent the influence that aerosols have on climate. On one hand the water content of aerosol
17 particles at atmospheric relative humidity (RH) below 100% contributes significantly to the
18 direct effect they have on the global radiative balance (Seinfeld and Pandis, 2006; Petters and
19 Kreidenweis, 2007; Swietlicki et al., 2008; Zieger et al., 2011; Rastak et al., 2014). On the other
20 hand the water-affinity of aerosol constituents, together with their dry size, defines the
21 efficiency with which these particles can activate as cloud condensation nuclei (CCN) at
22 supersaturated conditions (RH > 100%), form cloud droplets, and thus affect the properties of
23 clouds (Twomey 1974; Aalbrecht, 1989; McFiggans et al., 2006). To quantify the effects of
24 aerosol particles on clouds and climate it is thus necessary to understand the ways that aerosol
25 constituents interact with water.

26 Organic compounds contribute a large fraction (20-90%, depending on the environment)
27 of atmospheric submicron particulate mass (Jimenez et al., 2009) – which is the part of the
28 aerosol size distribution that typically dominates the CCN numbers. A significant fraction of
29 this organic aerosol (OA) is secondary – i.e. produced in the atmosphere from the condensation
30 of oxidation products of volatile, intermediate volatility and semi-volatile organic compounds
31 (VOCs, IVOCs and SVOCs). Emissions of biogenic VOCs such as monoterpenes, isoprene and

1 sesquiterpenes, followed by their subsequent oxidation and condensation in the atmosphere are
2 thought to be the dominant source of secondary organic aerosol (SOA) on a global scale
3 (Hallquist et al., 2009 and references therein) – although recent studies suggest also a notable
4 anthropogenic component to the global SOA (Volkamer et al., 2006; Hoyle et al., 2011;
5 Spracklen et al., 2011).

6 The solubility in water is one of the key properties governing the water-absorption (i.e.
7 hygroscopic growth) and CCN activation of aerosol particles. Together with aqueous phase
8 activity coefficients, surface tension, density and dry mass of the particle, water-solubility
9 affects the aerosol particle water content in thermodynamic equilibrium (Pruppacher and Klett,
10 1997; Seinfeld and Pandis, 2006; Topping et al., 2012). Atmospheric organic compounds have
11 a wide range of solubilities (Raymond and Pandis, 2003; Chan et al., 2008; Psichoudaki and
12 Pandis, 2013). OA is thus a complex mixture of molecules with different CCN-behaviour than
13 pure compounds. To accurately predict the water content and CCN activation of atmospheric
14 OA, information on the dissolution behaviour and aqueous phase interactions of these complex
15 mixtures is needed.

16 Representation of the complexity of OA is a major challenge for atmospheric chemical
17 transport models: OA consists of thousands of different compounds whose properties are poorly
18 known (Golstein and Galbally, 2007; Hallquist et al., 2009; Kroll et al., 2011). Approaches that
19 simplify the complex nature of the OA mixture, yet reproducing its behaviour accurately
20 enough, are required to be able to assess the climate and air quality effects of atmospheric
21 organics in large-scale modelling applications. One example of such approach is the
22 representation of the condensation and evaporation of SOA using a limited number of surrogate
23 compounds with a range of saturation concentrations, known as the volatility basis set (VBS,
24 Donahue et al., 2006; 2011; 2012). Similar simplifying approaches are needed to represent the
25 hygroscopic growth and CCN activation of OA as well.

26 When interpreting laboratory and field studies on hygroscopicity and CCN activation, a
27 number of simplifying assumptions about the OA properties have been made, for instance: 1)
28 assuming that organics completely dissolve in water at the point of activation (Huff-Hartz et
29 al., 2006); 2) assuming a fraction (ε_{eff}) of organics to be completely soluble and the remaining
30 fraction ($1 - \varepsilon_{eff}$) completely insoluble in water (e.g. Pruppacher and Klett, 1997; Engelhart et
31 al., 2008); 3) **lumping the phase-equilibrium thermodynamics, molar masses and densities of
32 the OA constituents into a single semi-empirical parameter. One of the most commonly-used**

1 formulations is the hygroscopicity parameter κ , which relates the water activity in the aqueous
2 solution with the water and dry particle volumes, and can be modified to account for limited
3 solubility as well if the solubilities of the individual aerosol constituents are known (Petters and
4 Kreidenweis, 2007; 2008; 2012, Petters et al., 2009a-c; Farmer et al., 2015; see also Rissler et
5 al., 2004 and Wex et al., 2007 for alternative single-parameter formulations). These common
6 simplifications of organic aerosol solubility and hygroscopicity are summarized in Table 1.

7 Laboratory studies on different types of organic aerosols have provided important
8 insights into the relationship between CCN-activation, hygroscopic growth and water-solubility
9 of the atmospheric OA constituents. Raymond and Pandis (2002, 2003) and Chan et al. (2008)
10 investigated the CCN activation of single- and multi-component aerosol particles consisting of
11 organic compounds with known solubilities in water, and found that the particles activated at
12 lower supersaturations than would have been expected based on the bulk solubility of their
13 constituents. As an example, the laboratory studies by Chan et al. (2008) indicate that the CCN
14 activation of material with water solubility as low as 1 g L^{-1} could be predicted assuming
15 complete dissolution. For some model systems the surface properties (wettability) of the aerosol
16 particles, instead of the bulk water-solubility, seemed a more important factor defining their
17 CCN activation (Raymond and Pandis, 2002). Huff Hartz et al. (2006) assigned part of this
18 effect to residual water left in the particles upon their generation, causing the particles to exist
19 as metastable aqueous solutions and thus activate at lower supersaturations than the
20 corresponding dry material. The rest of the apparent increase in solubility was assigned to
21 potential impurities in the particles. In general the results reported by Huff Hartz et al. (2006)
22 suggested that compounds with water-solubilities above 3 g L^{-1} behaved as if they were
23 completely soluble in water, in general agreement with the earlier results of Hori et al. (2003).

24 Secondary organic aerosol particles generated in the laboratory through oxidation
25 chemistry and condensation of the reaction products have also been found to activate as cloud
26 droplets and thus contribute to the atmospheric CCN budgets (Cruz and Pandis, 1997; 1998;
27 Huff Hartz et al., 2005; VanReken et al., 2005; Prenni et al., 2007; King et al., 2007; 2009;
28 Engelhart et al., 2008; 2011; Asa-Awuku, 2009; 2010). These particles probably resemble the
29 real atmospheric SOA more closely than individual organic species or their simple mixtures,
30 but the theoretical interpretation of their CCN-behaviour is complicated by the variety of their
31 constituents. Despite the fact that CCN-activity of SOA has been reported to vary with the
32 volatile precursor identity and loading (Varutbangkul et al., 2006; King et al., 2009; Good et

1 al., 2010), photochemical aging (Duplissy et al., 2008; Massoli et al., 2010, and temperature
2 (Asa-Awuku et al., 2009), the reported hygroscopicity parameter κ values determined for
3 different SOA types are remarkably similar, being typically around 0.1 for the overall SOA and
4 0.3 for the dissolved fraction extracted from the aqueous sample (Asa-Awuku et al., 2010; King
5 et al., 2010). Similarly, Huff Hartz et al. (2005) reported effective solubilities of as high as 100
6 g L⁻¹ for both mono- and sesquiterpene SOA – although both are known to consist of a range
7 of compounds with different solubilities. These results demonstrate the importance of knowing
8 the water-soluble fraction of SOA at varying conditions but suggest that its exact speciation is
9 probably not necessary for predictive understanding of the CCN activity of SOA particles (Asa-
10 Awuku et al., 2010; Engelhart et al., 2011). The κ values inferred from sub- or supersaturated
11 conditions for the same SOA mixtures, on the other hand, are not always consistent, the sub-
12 saturated κ values being typically lower than the super-saturated ones (Prenni et al., 2007;
13 Duplissy et al., 2008; Wex et al., 2009; Topping et al., 2012). Multiple possible reasons for this
14 have been presented in the literature, including incomplete dissolution of the aerosol
15 constituents at sub-saturated conditions (Petters et al., 2009a), surface tension effects (Good et
16 al., 2010), RH-driven effects on the reaction chemistry and thus the composition (solubility and
17 activity) of the formed SOA (Poulain et al., 2010), evaporation and condensation of semi-
18 volatile organic compounds (Topping et al., 2012), or the non-ideality of the mixtures being
19 more pronounced at the sub-saturated conditions (Kreidenweis et al., 2006; Petters et al., 2009a;
20 Good et al., 2010).

21 While the basic theory of cloud droplet activation for pure water-soluble compounds is
22 relatively well-established (Pruppacher and Klett, 1997; Asa-Awuku et al., 2007; Topping et
23 al., 2012; Farmer et al., 2015), and a number of theoretical and experimental studies on the
24 different aspects controlling the CCN-activation of SOA have been presented (see above,
25 McFiggans et al., 2006; Dusek et al., 2006 and references therein), only few of these studies
26 have investigated the implications of the water-solubilities of complex organic mixtures on
27 CCN activation. Understanding the relationship between the dissolution behaviour and CCN-
28 activation of complex organic mixtures is, however, needed to constrain the water-soluble
29 fraction of SOA in varying conditions as well as to systematically unravel the mechanisms
30 causing the apparent simplicity in the CCN-behaviour of complex organic mixtures.

31 In this work we introduce a framework for representing the mixture components with a
32 continuous distribution of solubilities, similar to the VBS (Donahue et al., 2006; 2011; 2012).

1 Using this framework in a theoretical model, we investigate the dissolution behaviour of
2 complex organic mixtures and their CCN activity, focusing on the impact of mixture solubility
3 on CCN-activation. In particular, we study the response of the CCN-activation to varying
4 solubility ranges, distribution shapes, and numbers of components in the mixture. Furthermore,
5 we compare the CCN-activation predictions using the simplified solubility representations
6 outlined above (complete dissolution, soluble fraction ε_{eff} , and hygroscopicity parameter κ
7 without including knowledge about the component solubilities) with the more detailed
8 description using the full solubility distributions, and study the relationship of the simplified
9 solubility parameters ε_{eff} and κ to the true mixture solubility distribution. Although the solubility
10 ranges and other thermodynamic properties of the mixture have been chosen to represent SOA,
11 many of the concepts and approaches introduced here can be applied to any particles consisting
12 of complex mixtures of organic compounds with varying water-solubilities. Finally, we discuss
13 the applicability of the introduced framework for describing the water-interactions of realistic
14 SOA mixtures and the relevant future directions.

15

16 2 Methods

17 2.1 Theoretical predictions of CCN activation of complex organic mixtures

18 Figure 1 schematically summarizes the model system considered in this study. We consider a
19 monodisperse population of spherical aerosol particles consisting of an internal mixture of
20 organic compounds. When exposed to water vapour, these particles grow reaching
21 thermodynamic equilibrium between the water vapour and the particle phase. The wet particle
22 is allowed to consist of maximum two phases: the insoluble organic phase and the aqueous
23 phase. The compositions of the organic and aqueous phases are determined on the one hand by
24 the equilibrium between the aqueous phase and the water vapour, and on the other hand by the
25 equilibrium of the aqueous phase with the organic insoluble phase. To isolate the effects of
26 solubility from organic volatility effects, we do not allow the organics to evaporate from the
27 droplet – i.e. we assume that the equilibrium vapour pressures of the organics is zero above the
28 droplet surface. Similarly, no condensation of organics from the gas phase to the particles is
29 allowed to take place. The validity of this assumption depends on the gas-phase concentrations
30 of the organic species as well as the atmospheric temperature during the cloud formation
31 process. Testing it at different atmospherically relevant conditions deserves some future

1 attention, accounting for the dynamics of the atmospheric gas-phase as well (see also Topping
2 et al., 2012). In this study, however, we focus strictly on the CCN-activation process. The
3 organic composition and dry particle size were treated as an input to a model calculating the
4 final equilibrium composition, wet size, and CCN activation behaviour of these particles. Note
5 that while the solubility in the equations presented in the next Sections 2.1 and 2.2 is non-
6 dimensional ($\text{g g}_{\text{H}_2\text{O}}^{-1}$), in the presentation of the results it is converted into g L^{-1} , assuming
7 constant unit density of water.

8 **2.1.1 Equilibrium between water vapour and an aqueous phase containing dissolved** 9 **material**

10 The Köhler equation (Pruppacher and Klett, 1997) is used to link the ambient water vapour
11 saturation ratio S with the size, composition and water content of the aerosol particles in
12 thermodynamic equilibrium (lower panel of Fig. 1):

$$13 \quad S = \frac{p_{w,eq}}{p_{w,sat}} = a_w \exp\left(\frac{4\sigma v_w}{RTD_{p,wet}}\right), \quad (1)$$

14 where $p_{w,eq}$ (Pa) is the equilibrium vapor pressure of water over the droplet surface, $p_{w,sat}$ (Pa)
15 the saturation vapor pressure over a pure flat water surface, σ (N m^{-1}) is the surface tension of
16 the droplet, v_w the molar volume of water in the aqueous phase, M_w (kg mol^{-1}) the molar mass
17 of water, ρ (kg m^{-3}) the density of the aqueous phase, $D_{p,wet}$ (m) the diameter of the droplet, T
18 (K) the temperature and R ($\text{J mol}^{-1} \text{K}^{-1}$) the universal gas constant. a_w is the water activity,
19 defined as the product of the water mole fraction X_w and water activity coefficient in the aqueous
20 phase Γ_w :

$$21 \quad a_w = X_w \Gamma_w. \quad (2)$$

22 The activity coefficient describes the interactions between water molecules and the dissolved
23 organic molecules in the mixture. The saturation ratio at which the particles of dry size $D_{p,dry}$
24 activate as cloud droplets (i.e. continue growing in size even if the saturation ratio decreases),
25 is referred to as the critical saturation ratio S_c . Mathematically this corresponds to the highest
26 local maximum in the $S(D_{p,wet})$ curve, usually referred to as the Köhler curve.

27 **2.1.2 Equilibrium between the aqueous and insoluble organic phases**

28 The composition of the droplet and the distribution of material between the organic insoluble
29 and the aqueous phases can be calculated applying the principles of mass conservation and the

1 thermodynamic equilibrium of the organic components in an aqueous mixture with the
 2 insoluble organic phase. As the mass transfer of organics between the particles and the gas
 3 phase is neglected, the total mass of the dry particle m_{dry} , being the sum over all components i ,
 4 is equal to the total organic mass in the wet droplet (see Fig. 1):

$$5 \quad m_{dry} = \sum_i^n m_{i,insoluble} + \sum_i^n m_{i,aqueous} , \quad (3)$$

6 where n is the total number of organic compounds, $m_{i,insoluble}$ is the mass of compound i in the
 7 insoluble organic phase and $m_{i,aqueous}$ the mass of compound i in the aqueous phase. The same
 8 holds for each organic compound individually:

$$9 \quad m_{i,dry} = y_{i,dry} m_{dry} = m_{i,insoluble} + m_{i,aqueous} , \quad (4)$$

10 where $y_{i,dry}$ is the mass fraction of i in the dry organic particle. On the other hand, the
 11 concentration of each organic compound in the aqueous phase is determined by the
 12 thermodynamics of the two-phase system consisting of the insoluble organic phase and the
 13 aqueous solution phase. The mass of each organic compound i in the aqueous phase can be
 14 expressed as (Prausnitz et al., 1998; Banerjee, 1984):

$$15 \quad m_{i,aqueous} = \begin{cases} \gamma_i Y_{i,wet} c_{sat,pure,i} m_w , & Y_{i,wet} > 0 \\ m_{i,dry} , & Y_{i,wet} = 0 \end{cases} \quad (5)$$

16 where γ_i is the activity coefficient of i in the insoluble organic phase (where the reference state
 17 is the pure component dissolution to water), $Y_{i,wet}$ and $c_{sat,pure,i}$ (here in $\text{g g}_{\text{H}_2\text{O}}^{-1}$) are the organic
 18 phase mole fraction and pure component solubility (saturation concentration) of i , and m_w is the
 19 total mass of water in the droplet. The former equation corresponds to the situation where the
 20 particle contains an insoluble organic core in thermodynamic equilibrium, the latter to the case
 21 where only the aqueous phase exists, i.e. all the organic material has dissolved to the water.
 22 Although the mole fraction and the corresponding molar activity coefficient have been used in
 23 Eq. 5, a similar relationship can be defined using the mass fraction in the organic phase and a
 24 corresponding mass-based activity coefficient. For a multicomponent system in which the
 25 molar mass of the organic species varies, the mole and mass fractions of a given species are not
 26 necessarily equal. In this study, however, we assume a constant molar mass throughout the
 27 organic mixture for simplicity, leading to the mass and mole fractions in the organic phase to
 28 be the same, i.e. $Y_i = y_i$ for all compounds. All the equations presented below can be re-derived

1 in a relatively straightforward manner taking into account a potential difference between the
 2 mole and the mass fractions in the organic phase.

3 Finding the organic and aqueous phase compositions that satisfy Eqs. 3-5 for given
 4 water and dry particles masses (m_w and m_{dry} , respectively) requires solving n coupled equations.
 5 These equations were expressed using the ratio χ_i of organic compound i in the insoluble core
 6 of the wet particle to the total mass of the compound (Raymond and Pandis, 2003; Petters and
 7 Kreidenweis 2008):

$$8 \quad \chi_i = \frac{m_{i,insoluble}}{m_{i,insoluble} + m_{i,aqueous}} = \frac{m_{i,insoluble}}{m_{i,dry}} = \frac{m_{i,insoluble}}{Y_{i,dry} m_{dry}}. \quad (6)$$

9 The mole fraction (equal to the mass fraction for the mixtures considered here) of i in the
 10 insoluble core is defined as

$$11 \quad Y_{i,wet} = \frac{m_{i,insoluble}}{\sum_i m_{i,insoluble}} = \frac{\chi_i m_{i,dry}}{\sum_i \chi_i m_{i,dry}} = \frac{\chi_i Y_{i,dry}}{\sum_i \chi_i Y_{i,dry}}. \quad (7)$$

12 Finally, combining equations 3-7, we get n equations of the form

$$13 \quad \chi_i = 1 - \frac{\chi_i Y_{i,dry} c_{i,sat,pure} m_w}{m_{i,dry} \sum_i \chi_i Y_{i,dry}}, \quad (8)$$

14 which can be solved for χ_i with the constraint $0 \leq \chi_i \leq 1$ for given water and dry particle masses.

15 **2.1.3 Representation of complex organic mixtures: Solubility distributions and** 16 **thermodynamic properties**

17 A novel aspect of this study as compared with previous theoretical work is the representation
 18 of complex mixtures using their aqueous solubility distribution of the individual species. In our
 19 calculations we used mixtures of n compounds, whose water-solubilities ranged from $c_{sat,min}$ to
 20 $c_{sat,max}$, either on a linear or logarithmic basis. The shape of the distribution could vary as well.
 21 In this work we studied essentially three types of mass fraction distributions in the dry particle:
 22 a uniform distribution in which all solubilities are equally abundant, distribution increasing
 23 steadily (linearly or logarithmically), and a distribution decreasing steadily (linearly or
 24 logarithmically). The 72 studied solubility distributions are specified in Table 2, and the
 25 solubility distributions for $n = 10$, $c_{sat,min} = 0.1 \text{ g L}^{-1}$ and $c_{sat,max} = 1000 \text{ g L}^{-1}$ are presented in
 26 Fig. 2 as examples.

1 For simplicity, we assumed that water forms an ideal solution with the dissolved
2 organics, i.e. $\Gamma_w = 1$, thus yielding an activity equal to the mole fraction of water, $a_w = X_w$ in Eq.
3 1. Since information about the activity coefficients of organic species in purely organic mixtures
4 is still scarce, we studied two alternative approaches to represent the dissolution
5 thermodynamics of the SOA mixture in Eqs. 9-12: 1) assuming an ideal **organic** mixture where
6 $\gamma = 1$ for all compounds **in the insoluble phase**; 2) assuming a constant organic phase activity
7 $\gamma_i Y_{i,wet}$ of unity for all compounds – in which case the dissolution behaviour of each i is similar
8 to their behaviour as pure components. These cases probably represent the limiting cases for
9 the dissolution of SOA components in CCN activation reasonably well, the former representing
10 a lower limit and the latter an upper limit for the overall solubility of the dry particle. Applying
11 the two limiting assumptions about the interactions of the compounds in the organic phase for
12 the 72 different solubility distributions (Table 2) thus results in total 144 unique representative
13 model mixtures.

14 The density, surface tension, and molar masses assumed for water and the organic
15 compounds are summarized in Table 3. Although the density, surface tension and molar mass
16 of the organics are likely to vary with the solubility, we kept them constant throughout the
17 organic mixture to isolate the solubility effects on the CCN behaviour. The values were chosen
18 based on literature studies of the CCN behaviour of SOA (Engelhart et al., 2008; Asa-Awuku
19 et al., 2010). The surface tension σ **was approximated by the surface tension of water, and the**
20 **molar volume of water in the aqueous phase was assumed the same as for pure water.**
21 Furthermore, we assumed no dissociation of the organics in the aqueous phase.

22 **2.1.4 Model calculations**

23 We solved Eq. 8 for organic mixtures with the Matlab internal function *fsolve*, for varying water
24 and dry particle masses m_w and m_{dry} , covering 50 different dry particle diameters between 20
25 and 500 nm. The calculations yielded the composition of the insoluble organic and the aqueous
26 phase, and thus the mole fraction of water in the aqueous solution X_w . From these results the
27 Köhler curves $S(D_{p,wet})$ corresponding to each dry particle mass could be calculated using Eq.
28 1 (see Fig. 3a for an example of the Köhler curves). The critical supersaturations s_c (defined as
29 $S_c - 1$) corresponding to specific dry particle diameters $D_{p,dry}$ (termed also as activation
30 diameters $D_{p,act}$ at a given saturation ratio S or supersaturation s) were determined from the
31 maxima of the Köhler curves (see Fig. 3a). The temperature was assumed to be 298 K in all
32 calculations. These calculations for the 144 unique organic model mixtures corresponded to

1 7200 Köhler curves yielding 5957 ($D_{p,act}$, s_c) pairs (activation points, see Fig. 3b). For the
2 remaining 1143 curves no activation points were found with the given combinations of mixture
3 properties and dry diameters. For comparisons with the simple solubility representations, the
4 dissolved organic fraction defined as

$$5 \quad \varepsilon = \frac{\sum m_{i,aqueous}}{m_{dry}} \quad (9)$$

6 was extracted from the model output.

7 **2.2 Comparison of the full model output to simple solubility representations**

8 To investigate the performance of the simple solubility representations given in Table 1 in
9 reproducing the CCN activation of complex mixtures, we fitted the ($D_{p,act}$, s_c) data created by
10 the full model using these simpler models. No fitting is required for the complete solubility
11 approach. Using the obtained solubility parameters from the optimal fit and the corresponding
12 simplified forms of the Köhler equation we then recalculated new ($D_{p,act}$, s_c) pairs and compared
13 them to the predictions by the full model. Furthermore, we investigated the relationships
14 between the true mixture solubility distribution and the simplified solubility parameters. The
15 details of the approach used for each simple model are outlined below.

16 **2.2.1 Complete dissolution**

17 In the case where all of the organic material is assumed to completely dissolve at the point of
18 activation, the calculation of the aqueous solution composition becomes trivial as

$$19 \quad m_{i,dry} = m_{i,aqueous} \quad (10)$$

20 for all the compounds, and the water mole fraction can simply be calculated based on the dry
21 particle mass as

$$22 \quad X_w = \frac{\frac{m_w}{M_w}}{\frac{m_w}{M_w} + \frac{m_{dry}}{M_{org}}}, \quad (11)$$

23 where m_w is the water mass in the droplet, m_{dry} the dry particle mass (related to $D_{p,dry}$ through
24 the organic density ρ_{org}) and M_{org} the organic molar mass. The X_w calculated in this way was
25 inserted into Eq. 1 to yield the corresponding ($D_{p,act}$, s_c) predictions and was also applied for

1 calculating the solution density and surface tension as mass-weighted averages of the water and
2 pure organic values.

3 **2.2.2 Hygroscopicity parameter κ**

4 In many practical applications the water activity and the difference in the densities and molar
5 masses of water and the dry material are expressed with a single hygroscopicity parameter κ ,
6 introduced by Petters and Kreidenweis (2007), defined as:

$$7 \quad \frac{1}{a_w} = 1 + \kappa \frac{V_s}{V_w}, \quad (12)$$

8 where V_s and V_w are the volumes of the dry material and water, respectively. The following
9 formulation of the relationship between water saturation ratio, aerosol size and composition is
10 referred to as the κ -Köhler equation:

$$11 \quad S = \frac{D_{p,wet}^3 - D_{p,dry}^3}{D_{p,wet}^3 - D_{p,dry}^3 (1 - \kappa)} \exp\left(\frac{4\sigma M_w}{RT\rho D_{p,wet}}\right) \quad (13)$$

12 yielding an approximate expression for the relationship between s_c and $D_{p,act}$ defined as

$$13 \quad s_c = \frac{2}{3} \left(\frac{4M_w \sigma}{RT\rho}\right)^{3/2} (3\kappa D_{p,act}^3)^{-1/2}, \quad (14)$$

14 Equation (14) was fitted to all ($D_{p,act}$, s_c) data produced for a given organic mixture composition
15 (see Table 2) by the full model, thus assuming a constant κ value for a given organic mixture.
16 To mimic the application of Eq. 14 to experimental data with no knowledge on the exact solute
17 composition, in this case we assumed the surface tension and density to be those of water when
18 fitting the κ values to the full model data. **The above formulation of κ , which is often used in
19 the interpretation of experimental data as well, thus contains information about solubility,
20 potential aqueous-phase non-ideality, as well as molar mass and density of the solutes (see
21 Farmer et al., 2015).**

22 **2.2.3 Soluble fraction ε_{eff}**

23 For an ideal solution of water and an organic solute the κ is directly proportional to the dissolved
24 fraction and the ratio of the molar volumes of water and the solute i.e.: $\kappa = \varepsilon \kappa_{max}$, where $\kappa_{max} =$
25 $(M_w/M_{org})(\rho_{org}/\rho_w)$. Assuming that a single soluble fraction ε_{eff} can represent a given organic

1 mixture (see Table 2) at all considered supersaturations, and substituting these relationships
 2 into Eq. 16 yields

$$3 \quad s_c = \frac{2}{3} \left(\frac{4M_w \sigma}{RT\rho} \right)^{3/2} \left(3 \frac{M_w}{M_s} \frac{\rho_s}{\rho_w} \varepsilon_{eff} D_{p,act}^3 \right)^{-1/2}, \quad (15)$$

4 the corresponding form of the Köhler equation being (see Huff Hartz et al., 2005)

$$5 \quad S = \frac{M_{org} \rho_w (D_{p,wet}^3 - D_{p,dry}^3)}{D_{p,wet}^3 (M_{org} \rho_w) + D_{p,dry}^3 (\varepsilon_{eff} M_w \rho_{org} - M_{org} \rho_w)} \exp\left(\frac{4\sigma M_w}{RT\rho D_{p,wet}} \right). \quad (16)$$

6 Again, we fitted Eq. 17 to the data produced by the full model and assumed the aqueous solution
 7 density and surface tension to be equal to those of water. When $\varepsilon_{eff} < 1$, the following
 8 relationship has been used to estimate the effective saturation concentration of the mixture
 9 (Raymond and Pandis, 2002; Huff Hartz et al., 2005)

$$10 \quad c_{sat,eff} = \frac{\rho_w (D_{p,wet}^3 - D_{p,act}^3)}{\varepsilon_{eff} \rho_{org} D_{p,act}^3}. \quad (17)$$

11 2.2.4 Connection between ε_{eff} and the solubility distribution of the mixture

12 Let us now assume that the dissolved fraction at the point of activation for each considered
 13 mixture can be expressed as a sum of two terms, the contribution from the compounds below a
 14 threshold solubility bin i_t and the contribution from the compounds over the threshold:

$$15 \quad \varepsilon = \frac{\sum_{i=1}^{i_t} m_{i,aqueous} + \sum_{j=i_t+1}^n m_{j,aqueous}}{m_{dry}} = \frac{\sum_{i=1}^{i_t} m_{i,aqueous} + \left(\sum_{j=i_t+1}^n m_{j,dry} - \sum_{j=i_t+1}^n m_{j,insoluble} \right)}{m_{dry}}. \quad (18)$$

16
 17 We now hypothesize that assuming a single soluble fraction for a given aerosol mixture is in
 18 fact equivalent to assuming that everything above i_t is completely dissolved while all the
 19 material below this threshold remain undissolved, i.e.

$$20 \quad \varepsilon_{eff} = \frac{\sum_{j=i_t+1}^n m_{j,dry}}{m_{dry}}. \quad (19)$$

21 On the other hand, $\varepsilon = \varepsilon_{eff}$ if the following condition is fulfilled (see Eq. 20):

$$\sum_{i=1}^{i_t} m_{i, \text{aqueous}} = \sum_{j=i_t+1}^n m_{j, \text{insoluble}} = \sum_{j=i_t+1}^n m_{j, \text{dry}} - \sum_{j=i_t+1}^n m_{j, \text{aqueous}} \quad (20)$$

Substituting Eq. 5 to Eq. 20 we now have

$$F_w \sum_{i=1}^{i_t} \gamma_i Y_{i, \text{wet}} C_{\text{sat}, \text{pure}, i} = \sum_{j=i_t+1}^n Y_{\text{dry}, j} - F_w \sum_{j=i_t+1}^n \gamma_j Y_{j, \text{wet}} C_{\text{sat}, \text{pure}, j} \quad (21)$$

where

$$F_w = \frac{m_w}{m_{\text{dry}}} \quad (22)$$

At the limit of large n and in the case of a symmetric distribution of material between the insoluble organic and aqueous phases, Eq. 21 is satisfied by setting the threshold solubility i_t so that

$$\lim_{i \rightarrow i_t} F_w \gamma_i Y_{i, \text{wet}} C_{\text{sat}, \text{pure}, i} = Y_{\text{dry}, i_t} - F_w \gamma_i Y_{i, \text{wet}} C_{\text{sat}, \text{pure}, i} \quad (23)$$

In this case the threshold solubility c_t is found from the bin for which

$$c_{\text{sat}, \text{pure}, i_t} = c_t \approx \frac{Y_{\text{dry}, i_t}}{Y_{\text{wet}, i_t}} \cdot \frac{1}{\gamma_{i_t}} \cdot \frac{1}{2F_w} \quad (24)$$

This is also equal to the bin where 50% of the material is partitioned in the insoluble phase, i.e.

$$\chi_{i_t} = \frac{m_{i_t, \text{insoluble}}}{m_{i_t, \text{insoluble}} + m_{i_t, \text{aqueous}}} = \frac{1}{1 + \frac{m_{i_t, \text{aqueous}}}{m_{i_t, \text{insoluble}}}} = \frac{1}{2} \quad (25)$$

Finding the solubility threshold c_t requires knowledge on the ratio F_w (Eq. 22). F_w , on the other hand, depends on the ambient supersaturation and the total soluble mass – thus introducing a supersaturation-dependence to the ε given by Eq. 18 as well. The magnitude of F_w as a function of supersaturation can be estimated by substituting Eqs. 15 and 17 into the definition of F_w (Eq. 22) and after some rearranging, yields:

$$F_w = \frac{\rho_w}{\rho_{\text{org}}} \left(\frac{2\varepsilon}{s_c} \cdot \frac{\rho_{\text{org}}}{\rho_w} \cdot \frac{M_w}{M_{\text{org}}} - 1 \right) = \frac{\rho_w}{\rho_{\text{org}}} \left(\frac{2\varepsilon \kappa_{\text{max}}}{s_c} - 1 \right) \quad (26)$$

20

1 3 Results

2 Figure 3a displays examples of the Köhler curves obtained from solving Eqs. 1 and 6 for
3 distribution 1 (the flat logarithmic distribution) with varying solubilities and $n = 5$, assuming
4 that the organics form an ideal mixture with each other. Each curve corresponds to a different
5 dry size, and the dots indicate the activation point (S_{crit} corresponding to the activation dry
6 diameter $D_{p,act}$). Black dots indicate incomplete dissolution ($\epsilon < 0.99$) at the point of activation
7 while red dots indicate that in practice all the organics are dissolved into the aqueous phase at
8 the point of activation ($\epsilon > 0.99$). Qualitatively similar behaviour was observed for all the
9 considered distributions: as the overall solubility of the mixture increases, the dissolution of the
10 compounds increases, leading eventually to complete dissolution at the point of activation. The
11 transition from a regime with two phases (aqueous + insoluble) to a single aqueous phase is
12 visible in the two maxima in Fig. 3a, in accordance with Shulman et al. (1996) and Petters and
13 Kreidenweis (2008).

14 Figure 3b illustrates the parameter space probed in this study, showing the 5957 (S_{crit} ,
15 $D_{p,act}$) points corresponding to the Köhler curve maxima calculated for all the considered
16 organic mixtures. The relationships between the critical supersaturation, activation diameter,
17 and dissolved fraction ϵ at the point of activation are also schematically shown. The chosen dry
18 diameters and supersaturations represent a conservative range of typical atmospheric conditions
19 – as the total aerosol number concentrations are dominated by ultrafine (diameters smaller than
20 100 nm) particles at most locations. In most considered cases the dissolved fractions fall
21 between 0.1 and 1, but the lowest dissolved fractions at the point of activation are of the order
22 of only a few percent – thus mimicking nearly insoluble aerosols. Therefore the cases
23 considered here represent a reasonable sample of atmospherically relevant conditions and SOA
24 mixture compositions. The water-to-organic mass ratios F_w corresponding to the probed
25 conditions and mixtures (see Sect. 2.2.4) range from values below 1 up to 1000, with most
26 values around 10-100. In many of the following plots and considerations we have chosen four
27 specific supersaturations, 0.1%, 0.3%, 0.6% and 1%, as representative values for typical
28 laboratory experiments, which are also indicated in Fig. 3b.

29 An example of the dependence of the activation diameter $D_{p,act}$ on the solubility range
30 for all the studied distributions (see Table 2) and $n = 5$ is presented in Fig. 4. As expected, the
31 activation diameter decreases with increasing supersaturation and solubility range for a given
32 solubility distribution. The solubility distribution is reflected in the overall magnitude of the

1 activation diameters: the distributions that have larger fractions of material in the higher end of
2 the solubility range (distributions 2, 3, and 4) have generally lower activation diameters for a
3 given supersaturation as compared with the other distributions. The case when unity activity in
4 the organic phase is assumed results in smaller activation diameters for the same supersaturation
5 as compared with the ideal organic mixture case (see Sect. 2.1.3 for the definitions of the cases).

6 Figure 5 presents the activation diameters predicted using the simplified solubility
7 descriptions (Table 1) based on best fits to all available data as compared with the full
8 description of the solubility distributions (Table 2). The results clearly show, not surprisingly,
9 that assuming complete dissolution for all the mixtures consistently under-predicts the
10 activation diameters (Fig. 5a). Representing the dissolution behaviour with only one additional
11 parameter, i.e. the hygroscopicity parameter κ (Eqs. 15-16) or the effective soluble fraction ε_{eff}
12 (Eqs. 17-18) improves the agreement between the activation diameters considerably (Figs. 5b
13 and 5c). Adding the knowledge about the molar mass and density of the organic mixture, which
14 is the only difference between using the ε_{eff} instead of the single κ , adds only marginal
15 improvements in predicting the activation diameters for a given supersaturation. The
16 disagreements between the simplified models and the full theoretical treatment are largest for
17 the smallest supersaturations. These are the cases with the widest range of possible ε values at
18 the point of activation (see Fig. 3b), and the effect is most obvious for the complete dissolution
19 model: the larger the deviation from complete dissolution at the point of activation (i.e. the $\varepsilon \sim$
20 1 case in Fig. 3b) the more significant error we introduce. The activation diameters predicted
21 assuming complete dissolution are within 10% of the correct values if the real dissolved fraction
22 ε is larger than about 0.7-0.8 at the point of activation.

23 The performance of the simple solubility models for all the studied Köhler curves is
24 summarized in Fig. 6: while the complete dissolution assumption results in systematic under-
25 prediction (up to 40%) of the activation diameter, the κ - and ε_{eff} -based models are generally
26 within 10% (in most cases within 5%) of the activation diameter predicted for the full solubility
27 distribution representation.

28 Figure 7 compares the fitted parameters representing the mixture dissolution to the
29 corresponding values inferred from the full mixture data for the 144 different mixtures. In Fig.
30 7a, the effective soluble fractions ε_{eff} are compared to the actual dissolved fractions ε at the point
31 of activation for all the studied mixtures. While in the fits a single constant ε_{eff} has been assumed
32 to represent a given mixture (see Eq. 15-16), in reality ε varies with supersaturation (Eq. 9).

1 Thus, while the fitted ε_{eff} for a given mixture correlates very well with the average ε over all
2 activation points (the markers in Fig. 7a), the performance of the approach can vary
3 considerably with supersaturation (the grey lines in Fig. 7a). In practice this means that
4 describing a given complex mixture with a fixed soluble fraction yields representative average
5 dissolution behaviour, but does not guarantee correct solubility description for a specific s_c if
6 fitted over a range of supersaturations. The corresponding comparison for the hygroscopicity
7 parameter κ values describing the data are shown in Fig. 7b. A clear correlation between the
8 fitted κ and the average ε is observed as expected (see Sect. 2.2.3), but the variation of ε with
9 supersaturation again adds scatter to the data – suggesting a dependence of κ on s_c . The
10 maximum κ , on the other hand, is defined primarily by the molar masses and densities of the
11 organics. For our mixtures with constant M_{org} and ρ_{org} the value of κ_{max} is 0.15, which is
12 indicated in Fig. 7b. The points above this theoretical maximum are a result of using the pure
13 water density instead of the mixture value in the Kelvin term of the Köhler equation (Eq. 16).
14 These results thus suggest that the κ values of 0.1-0.2 typically observed for SOA particles
15 (Duplissy et al., 2011) are controlled by the molar masses and densities of the SOA mixtures to
16 a large extent and can result from quite different SOA mixtures in terms of their solubilities.

17 To illustrate the relationship of the fitted ε_{eff} to the dissolution of a given mixture, the
18 partitioning between the aqueous and insoluble organic phase is presented in Fig. 8 for
19 distribution 1 with the “low” solubility range and $n = 100$ at the point of activation when $s_c =$
20 0.1% (see Table 2). Figure 8a shows the partitioning for the case where ideal organic mixture
21 has been assumed and Fig. 8b shows the corresponding data for the unity activity case (see Eq.
22 5). The point of 50%-partitioning (c_t , Eqs. 20-21) is also shown. As described in Sect. 2.2.4 we
23 expect c_t to be a reasonable estimate for the limit for complete dissolution, if the complex
24 mixture is reduced into a two-component mixture of completely soluble and insoluble
25 components. It should be noted, however, that the water content and ε of the droplet at the point
26 of activation depend on supersaturation (see Eq. 25), causing also a dependence between c_t and
27 s_c . Furthermore, Fig. 8 illustrates a difference in the solubility-dependence of the partitioning
28 behavior for the two organic activity assumptions. The ideal mixture displays a symmetrical
29 sigmoidal dependence around c_t . For the unity activity case, on the other hand, the undissolved
30 fraction is asymmetric around the 50%-value – dropping rapidly to zero above c_t but
31 approaching 1 asymptotically below c_t .

1 Figure 9 shows the distributions of the solubility bins containing the 50%-partitioning
2 points (c_t , Eqs. 20-21) on a decadal basis for all the activation points studied, illustrating also
3 the differences between the two assumptions about the organic phase activity. The c_t values for
4 the ideal organic mixture (Fig. 9a, based on 2465 points) display a symmetrical distribution
5 around the median value of about 10 g L^{-1} . Also, a modest dependence of c_t on the number of
6 components is observed: the cases with 3 and 5 components display slightly higher c_t values as
7 compared with the cases with larger n . This apparent dependence is probably due to the discrete
8 nature of the solubility distributions in combination with the fact that for the different solubility
9 ranges (see Table 2) only the lower end of the distribution is changed while the upper end is
10 always at 1000 g L^{-1} . The unity activity case displays a much stronger n -dependence (Fig. 9b,
11 based on 3492 points): if analysed separately, the median c_t shifts from about 0.1 to 10 g L^{-1}
12 when n changes from 100 to 10 and 5, and up to 100 g L^{-1} for $n = 3$. Unlike the ideal mixture
13 this behavior is explained by the actual dissolution thermodynamics: in a system where the
14 components do not affect each other's solubility directly (i.e. the dissolution of a compound i
15 is only limited by its own presence in the aqueous phase), the amount of dissolved material is
16 only dependent on the total water content and is larger the larger the number of dissolvable
17 components. If all the different mixtures are integrated together the median c_t for the unity
18 activity assumption lies at about 1 g L^{-1} – a decade lower than for the ideal mixture case. **Figure**
19 **10 provides a more detailed look on the sensitivity of the c_t to supersaturation, n and molar mass**
20 **for one of the distributions (Distribution 1, mid solubility range, see Table 2). As expected, c_t**
21 **depends considerably on supersaturation. It can also be seen that while c_t shows some sensitivity**
22 **to the number of components (in line with Fig. 9) and molar mass, in our case by far the most**
23 **critical assumption is related to the organic mixture thermodynamics.**

24 The fitted ε_{eff} values are compared in Figure 11 to the fraction of mass with solubilities
25 above the median c_t for the 134 mixtures that activated at the probed conditions, individually
26 for both organic activity assumptions. The fitted dissolved fraction corresponds well to the
27 fraction of mass with solubilities above the 50%-partitioning point, as predicted by the
28 theoretical principles outlined in Sect. 2.2.4. Also, the solubilities of the different distributions
29 with varying shapes, numbers of components and solubility ranges can be represented
30 reasonably well with a single median c_t (equal to 10 g L^{-1} for the ideal mixture case and 1 g L^{-1}
31 for the unity activity case, see Fig. 9) with median deviations between the fitted ε_{eff} and the
32 fraction above c_t of 9% and 8%, respectively. On the other hand, these results indicate that the
33 soluble fractions determined from experimental data on CCN activation provide information

1 about the fraction of material with solubility above c_t . There are 8 points that do not seem to
2 follow the general trend, however: a group of points with all the material above the threshold
3 solubility can display a variety of fitted ε_{eff} values. These are all points that correspond to the
4 “high” solubility ranges. Distributions 5 and 6 (see Table 2) with $n = 3$ are among these points
5 for both organic activity assumptions. For the ideal mixture case also distribution 1 with 3
6 components diverges from the general trend, and for the unity activity assumption distribution
7 5 falls into the category regardless of the number of components. These points thus contribute
8 to the high ends of the c_t distributions in Fig. 9.

9 The deviations in the activation diameters as predicted by the three simplified solubility
10 representations (complete dissolution, ε and κ) are displayed in Figure 12 as a function of the
11 mixture properties for $s_c = 0.1\%$, 0.3% , 0.6% and 1% . Again, the two different assumptions
12 about the organic phase activity are treated separately due to their different limiting solubilities
13 c_t (Fig. 9) and different shapes of the partitioning distributions (Fig. 8). Also, the points close
14 to complete dissolution at the point of activation ($\varepsilon \geq 0.8$, see Fig. 5 and its explanation in the
15 text) are presented with a different color (gray symbols) than the points where the activation
16 diameter differs significantly from the complete dissolution prediction ($\varepsilon < 0.8$, black symbols).
17 As expected, the complete dissolution assumption performs better for the more water soluble
18 organic mixtures. Figs. 12a and 12b illustrate this by showing the relationship between the norm
19 of the error in the predicted $D_{p,act}$ and the fraction of material below the median c_t (10 and 1 g
20 L^{-1}) for the two organic activity assumptions. The larger the amount of material below the
21 solubility limit, the larger the deviation from the full model predictions. For the κ and ε models,
22 on the other hand, the variable best correlating with the error in $D_{p,act}$ induced by the
23 simplification is different for the different organic activity assumptions – although close to
24 complete dissolution these models also do well, nearly independent of the solubility of the
25 distribution. For the ideal mixture case the fraction of mass between 1 and 100 g L^{-1} correlates
26 better with the error (Fig. 12c, 12e) than the mass fraction below any solubility limit (not
27 shown), while for the unity activity case the material at the low end of the distribution (mass
28 fraction below 1 g L^{-1}) performs better (see Fig. 12d, f). The reason for this lies in the different
29 shapes of the partitioning distributions resulting from the two assumptions (Fig. 8). For the
30 symmetric partitioning curve of the ideal mixture case, the predicted ε and κ are most sensitive
31 to differences in the partitioning behaviour between compounds within the range of c_t
32 corresponding to the supersaturation and particle diameter ranges studied here, i.e. 1-100 g L^{-1}
33 (see Figs. 12c, 12e). Anything outside these boundaries will behave as completely soluble or

1 insoluble throughout the studied supersaturation space, thus not introducing a significant error
2 when constant ε is assumed to describe the mixture. However, the more the material that can
3 behave as either insoluble or soluble depending on the conditions, the larger error we introduce
4 by assuming a constant ε for a given mixture at any conditions. The story is different for the
5 unity activity case (Figs. 12d, 12f): as the shape of the partitioning distribution (Fig. 8) does not
6 depend on c_t , the compounds with solubilities below c_t will contribute relatively much more the
7 fitted ε_{ff} than for the previous case, and thus the more material there is in the “tail” of the
8 partitioning distribution, the worse the assumption about a single ε for the whole distribution.

9 To relate the theoretical work conducted here to realistic atmospheric organic aerosol
10 mixtures, Fig. 13 displays an example of a solubility distribution representing SOA formed
11 from dark ozonolysis of α -pinene (Chen et al., 2008). The solubilities have been estimated with
12 SPARC (see Wania et al., 2014 and references therein). The average molar masses and O:C
13 ratios in each solubility bin are also displayed, along with the c_t values corresponding to the
14 activation points with limited solubility – assuming that the organics form an ideal mixture with
15 each other. Most of the material is predicted to have solubilities between 1 and 100 g L⁻¹,
16 indicating that this fresh α -pinene SOA is at the critical range of solubilities for limited
17 dissolution at the point of activation. This in turn suggests that the observed difference between
18 the κ values inferred from hygroscopicity and CCN-activation for this mixture might largely
19 result simply from the distribution of solubilities present.

20

21 4 Discussion and conclusions

22 We have studied the relationship between CCN activation and solubility of 144 different
23 theoretically constructed complex organic mixtures using Köhler theory, accounting for the
24 partial solubility of the compounds in water and assuming ideal interactions between the
25 dissolved molecules and water. The mixtures encompassed a wide variety of solubilities, and
26 were represented by solubility distributions with various solubility ranges and shapes
27 (analogously to the volatility basis set, VBS). Two limiting assumptions (ideal mixture vs. unity
28 activity) about the interactions between the organics in the insoluble organic phase were tested.
29 The results using this comprehensive solubility representation (termed as “the full model”) were
30 compared to commonly-used simplified descriptions of solubility: 1) assuming complete
31 dissolution; 2) representing the mixture with single hygroscopicity parameter κ ; 3) representing

1 the mixture with a single soluble fraction ε_{eff} . The calculations were carried out for particle dry
2 sizes ranging from 20 to 500 nm and supersaturations between 0.03% and 5%, thus probing an
3 atmospherically representative parameter space and resulting in total 5957 unique activation
4 points.

5 Comparing the full model predictions to the simplified solubility descriptions, we find
6 that assuming complete dissolution under-predicts the activation diameter up to about a factor
7 of two for the studied mixtures. Our results indicate that about 70-80% of the material needs be
8 dissolved at the point of activation for the complete dissolution assumption to predict activation
9 diameters that are within 10% of that produced by the full solubility treatment. Adding a single
10 parameter to describe the mixture solubility improved the situation considerably: the
11 predictions of activation diameters based on a single ε or κ for a given mixture were within 10%
12 of the full model predictions, the difference between these two approaches being only marginal.

13 The fitted soluble fractions, ε_{eff} , describing the solubility distribution (and thus the fitted
14 κ which is directly proportional to ε) were found to correspond well to the fraction of dry particle
15 material with solubilities larger than a given threshold solubility c_t . For the ideal organic
16 mixture assumption the median c_t was 10 g L⁻¹, most of the values falling between 1 and 100 g
17 L⁻¹, depending somewhat on the supersaturation. Since the material with solubilities outside
18 this range can generally be treated as completely soluble or insoluble in CCN activation
19 calculations, the error made by using the single soluble fraction increased when a larger fraction
20 of material was present in this critical range. For the unity activity case the median c_t was 1 g
21 L⁻¹, but decreased with the number of components present in the mixture, n . For the range of n
22 = 3-100 studied here, the typical c_t values were between 0.1 and 10 g L⁻¹. Due to the asymmetric
23 shape of the aqueous-organic phase partitioning of the organics in the unity activity case, the
24 simplified models performed better, the more material with solubilities larger than c_t was
25 present in the particles. In general, the median values for c_t represented the soluble fraction with
26 a reasonable accuracy in most of the studied mixtures, although the exact composition of the
27 mixtures varied considerably.

28 Our values for the limiting solubilities for complete dissolution are in agreement with
29 the values of 3 g L⁻¹ and 1 g L⁻¹ previously reported by Huff Hartz et al. (2006) and Chan et al.
30 (2008) based on experimental data on specific mixtures. Our results on the two different
31 assumptions about the organic phase activities indicate that the mixtures investigated by these
32 past studies were probably somewhat non-ideal, where the compounds hindered each other's

1 dissolution less than would be expected for a fully ideal mixture. On the other hand, in light of
2 our findings, the observations of the close to complete dissolution of SOA at activation (Huff
3 Hartz et al., 2005; Engelhart et al., 2011) indicate that the majority of the material in the studied
4 SOA mixtures had solubilities larger than 10 g L^{-1} . Our results suggest that even with vastly
5 different solubility distributions one can yield very similar CCN-activation behavior (and
6 consequently values of κ or ϵ), as the parameter that matters is the material above c_i .

7 The above results suggest that the solubility range corresponding to limited solubility in
8 CCN activation is between $0.1\text{-}100 \text{ g L}^{-1}$, and resolving the solubility distributions of aerosol
9 mixtures outside this range provides little added value for understanding their CCN activation.
10 In fact, this is probably a conservative estimate, as in most cases most material below 1 g L^{-1} is
11 practically insoluble and most material above 10 g L^{-1} completely soluble – even considering
12 the uncertainty in the organic mixture activity. These results can be used to guide the
13 representation of the cloud activation properties of complex mixtures, and provide quantitative
14 support for the previous notion that knowing the water-soluble fraction of the aerosol mixture
15 in question is the key in most applications. We provide quantitative estimates on how this
16 soluble fraction should be defined in the case of complex mixtures, and when such a simplified
17 model is not expected to perform well.

18 There are, however, some limitations to our approach to keep in mind when applying
19 the results to laboratory experiments or atmospheric data. Since the focus of this work was
20 strictly on the links between solubility and CCN activation, we did not explore in depth how
21 variation of surface activity, molecular mass, pure-component density, the gas-droplet
22 partitioning of the organic compounds or non-ideality of water with respect to the aqueous
23 phase would affect the results (see also Suda et al. 2012, 2014; Topping et al., 2012).
24 Furthermore, temperature was assumed to stay constant at 298 K. Since many of the
25 thermodynamic properties relevant to CCN activation are temperature-dependent (see e.g.
26 Christensen et al., 2012), future work investigating the impact of temperature on the phenomena
27 studied here is needed. Furthermore, the solubility and organic phase activity should naturally
28 be linked to the aqueous phase activity coefficients predicted and used in a number of previous
29 studies (see e.g. Topping et al., 2013 and references therein), although lack of well-defined
30 experimental data on organic phase activities and mixture solubilities currently hinders
31 quantitative evaluation of the current multi-phase mixture thermodynamic models (see e.g.
32 Cappa et al., 2008). Evaluation of the concepts and approaches presented here (e.g. the

1 solubility limits c_i) with laboratory studies on well-defined complex mixtures over a wide range
2 of solubilities, supersaturations and particle diameters would therefore be warranted.

3 Another important future step would be applying the introduced solubility distribution
4 framework in the atmospheric context. On one hand, the framework is likely to be useful in
5 modeling the evolution of the CCN-activity of secondary organics. We expect the solubility
6 distributions (and thus c_i) to depend on the SOA mixture properties such as the O:C ratio and
7 the molar masses of the mixture constituents, which in turn evolve due to atmospheric chemistry
8 (Kuwata et al., 2012; Suda et al., 2014), coupling the solubility distributions to the different
9 dimensions of the VBS (Donahue et al., 2006; Donahue et al., 2011; Kroll et al., 2011; Donahue
10 et al., 2012; Shiraiwa et al., 2014). A systematic study investigating the interlinkages between
11 these variables in light of the available experimental data from field and laboratory would thus
12 be a valuable future contribution. Furthermore, as atmospheric aerosol particles are typically
13 mixtures of organic and inorganic constituents, the molecular interactions between atmospheric
14 organics and inorganics as well as their effect of the pure-component solubility should be
15 expanded. On the other hand, the solubility distribution can be used as a simplifying concept
16 aiding in large-scale model simulations coupling atmospheric chemistry with the dynamics of
17 cloud formation. With the assumptions applied here the CCN-activation calculation itself is
18 computationally relatively light, slowing down considerably only if n is of the order 10000 or
19 larger with typical present-day computational resources. Therefore, using the solubility
20 distribution framework within an atmospheric model is a good option if accuracy beyond the
21 simple one-parameter approaches is required, or as an intermediate tool linking atmospheric
22 age to the effective ε or κ describing a given mixture.

24 Acknowledgements

25 Financial support from the European Research Council project ATMOGAIN (grant 278277),
26 Vetenskapsrådet (grant 2011-5120), the European Commission FP7 integrated project
27 PEGASOS (grant 265148) are gratefully acknowledged.

29 References

30 Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, *Science*, 245, 1227–
31 1230, 1989.

1 Asa-Awuku, A., and Nenes, A.: Effect of solute dissolution kinetics on cloud droplet formation:
2 Extended Köhler theory, *Geophys. Res.*, 112, D22201, doi: 10.1029/2005JD006934, 2007.

3 Asa-Awuku, A., Engelhart, G. J., Lee, B. H., Pandis, S. N., and Nenes, A.: Relating CCN
4 activity, volatility, and droplet growth kinetics of β -caryophyllene secondary organic aerosol,
5 *Atmos. Chem. Phys.*, 9, 795–812, 2009.

6 Asa-Awuku, A., Nenes, A., Gao, S., Flagan, R. C and Seinfeld, J. H.: Water-soluble SOA from
7 alkene ozonolysis: composition and droplet activation kinetics inferences from analysis of CCN
8 activity, *Atmos. Chem. Phys.*, 10, 1585–1597, 2010.

9 Banerjee, S.: Solubility of organic mixtures in water, *Environ. Sci. Technol.*, 18, 587–591,
10 1984.

11 Cappa, C. D., Lovejoy, E. R., and Ravishankara, A. R.: Evidence for liquid-like and nonideal
12 behavior of a mixture of organic aerosol components, *Proc. Natl. Acad. Sci. U. S. A.* 105,
13 18687, 2008.

14 Chan, M. N., Kreidenweis, S. M., and Chan, C. K.: Measurements of the hygroscopic and
15 deliquescence properties of organic compounds of different solubilities in water and their
16 relationship with cloud condensation nuclei activities, *Environ. Sci. Technol.* 42 3602–3608,
17 2008.

18 Chen, Q., Liu, Y., Donahue, N. M., Shilling, J. E., and Martin, S. T.: Particle-phase chemistry
19 of secondary organic material: Modeled compared to measured O:C and H:C elemental ratios
20 provide constraints, *Environ. Sci. Technol.*, 45, 4763, doi:10.1021/es104398s, 2011.

21 Christensen, S. I. and Petters, M. D.: The role of temperature in cloud droplet activation., *J.*
22 *Phys. Chem. A*, 116, 9706–17, doi:10.1021/jp3064454, 2012.

23 Cruz, C. N. and Pandis, S. N.: A study of the ability of pure secondary organic aerosol to act as
24 cloud condensation nuclei, *Atmos. Environ.*, 31, 2205–2214, doi:10.1016/S1352-
25 2310(97)00054-X, 1997.

26 Cruz, C. N. and Pandis, S. N.: The effect of organic coatings on the cloud condensation nuclei
27 activation of inorganic atmospheric aerosol, *J. Geophys. Res.*, 103, 13 111–13 123, 1998.

28 Donahue, N. M., Epstein, S. a., Pandis, S. N. and Robinson, a. L.: A two-dimensional volatility
29 basis set: 1. organic-aerosol mixing thermodynamics, *Atmos. Chem. Phys.*, 11, 3303–3318,
30 doi:10.5194/acp-11-3303-2011, 2011.

1 Donahue, N. M., Kroll, J. H., Pandis, S. N. and Robinson, a. L.: A two-dimensional volatility
2 basis set – Part 2: Diagnostics of organic-aerosol evolution, *Atmos. Chem. Phys.*, 12(2), 615–
3 634, doi:10.5194/acp-12-615-2012, 2012.

4 Donahue, N. M., Robinson, a L., Stanier, C. O. and Pandis, S. N.: Coupled partitioning, dilution,
5 and chemical aging of semivolatile organics., *Environ. Sci. Technol.*, 40, 2635–43 [online]
6 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16683603>, 2006.

7 Duplissy, J., DeCarlo, P. F., Dommen, J., Alfarra, M. R., Metzger, A., Barmapadimos, I., Prevot,
8 a. S. H., Weingartner, E., Tritscher, T., Gysel, M., Aiken, a. C., Jimenez, J. L., Canagaratna, M.
9 R., Worsnop, D. R., Collins, D. R., Tomlinson, J. and Baltensperger, U.: Relating
10 hygroscopicity and composition of organic aerosol particulate matter, *Atmos. Chem. Phys.*, 11,
11 1155–1165, doi:10.5194/acp-11-1155-2011, 2011.

12 Duplissy, J., Gysel, M., Alfarra, M. R., Dommen, J., Metzger, A., Prevot, S. H., Weingartner,
13 E., Laaksonen, A., Raatikainen, T., Good, N., Turner, S. F., McFiggans, G., and Baltensperger,
14 U.: Cloud forming potential of secondary organic aerosol under near atmospheric conditions,
15 *Geophys. Res. Lett.*, 35, L03818, doi:10.1029/2007GL031075, 2008.

16 Dusek, U., Frank, G. P., Hildebrandt, L., Curtius, J., Schneider, J., Walter, S., Chand, D.,
17 Drewnick, F., Hings, S., Jung, D., Borrmann, S., and Andreae, M. O.: Size matters more than
18 chemistry for cloud-nucleating ability of aerosol particles, *Science*, 312, 1375–1378, 2006.

19 Engelhart, G. J., Hildebrandt, L., Kostenidou, E., Mihalopoulos, N., Donahue, N. M., and
20 Pandis, S. N.: Water content of aged aerosol, *Atmos. Chem. Phys.*, 11, 911–920, 2011.

21 Engelhart, G. J., Moore, R. H., Nenes, A., and Pandis, S. N.: Cloud condensation nuclei activity
22 of isoprene secondary organic aerosol, *J. Geophys. Res.*, 116, D02207,
23 doi:10.1029/2010JD014706, 2011.

24 Farmer, D. K., Cappa, C. D. and Kreidenmweis, S. M.: Atmospheric processes and their
25 controlling influence on cloud condensation nuclei activity, *Chem. Rev.*, doi:
26 10.1021/cr5006292, 2015.

27 Goldstein, A. H. and Galbally, I. E.: Known and unexplored organic constituents in the earth's
28 atmosphere, *Environ. Sci. Technol.*, 41, 1514–1521, 2007.

29 Good, N., Topping, D. O., Duplissy, J., Gysel, M., Meyer, N. K., Metzger, a., Turner, S. F.,
30 Baltensperger, U., Ristovski, Z., Weingartner, E., Coe, H. and McFiggans, G.: Widening the

1 gap between measurement and modelling of secondary organic aerosol properties, *Atmos.*
2 *Chem. Phys.*, 10, 2577–2593, doi:10.5194/acp-10-2577-2010, 2010.

3 Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M. and
4 Dommen, J.: The formation, properties and impact of secondary organic aerosol: current and
5 emerging issues, *Atmos. Chem. Phys.*, 5155–5236, 2009.

6 Hori, M., Ohta, S., Muraio, N., and Yamagata, S.: Activation capability of water soluble organic
7 substances as CCN, *Aerosol Sci.*, 34, 219, 2003.

8 Hoyle, C. R., Boy, M., Donahue, N. M., Fry, J. L., Glasius, M., Guenther, A., Hallar, A. G.,
9 Huff Hartz, K., Petters, M. D., Petäjä, T., Rosenoern, T. and Sullivan, A. P.: A review of the
10 anthropogenic influence on biogenic secondary organic aerosol, *Atmos. Chem. Phys.*, 11, 321–
11 343, doi:10.5194/acp-11-321-2011, 2011.

12 Huff Hartz, K. E., Rosenoern, T., Ferchak, S. R., Raymond, T. M., Bilde, M., Donahue, N. M.,
13 and Pandis, S. N.: Cloud condensation nuclei activation of monoterpene and sesquiterpene
14 secondary organic aerosol, *J. Geophys. Res.*, 110, D14208, doi:10.1029/2004JD005754, 2005.

15 Huff-Hartz, K. E. H., Tischuk, J. E., Chan, M. N., Chan, C. K., Donahue, N. M., and Pandis, S.
16 N.: Cloud condensation nuclei activation of limited solubility organic aerosol, *Atmos. Environ.*,
17 40, 605–617, 2006.

18 Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, a S. H., Zhang, Q., Kroll, J. H.,
19 DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, a C., Docherty, K. S., Ulbrich, I. M.,
20 Grieshop, a P., Robinson, a L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. a, Hueglin,
21 C., Sun, Y. L., Tian, J., Laaksonen, a, Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M.,
22 Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. a,
23 Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick,
24 F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, a,
25 Miyoshi, T., Hatakeyama, S., Shimono, a, Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J.
26 R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, a M., Williams, L. R., Wood, E. C.,
27 Middlebrook, a M., Kolb, C. E., Baltensperger, U. and Worsnop, D. R.: Evolution of organic
28 aerosols in the atmosphere., *Science*, 326, 1525–1529, doi:10.1126/science.1180353, 2009.

29 King, S. M., Rosenoern, T., Shilling, J. E., Chen, Q. and Martin, S. T.: Cloud condensation
30 nucleus activity of secondary organic aerosol particles mixed with sulfate, *Geophys. Res. Lett.*,
31 34, L24806, doi:10.1029/2007GL030390, 2007.

1 King, S. M., Rosenoern, T., Shilling, J. E., Chen, Q., Wang, Z., Biskos, G., McKinney, K. A.,
2 Pöschl, U and Martin, S. T.: Cloud droplet activation of mixed organic-sulfate particles
3 produced by the photooxidation of isoprene, *Atmos. Chem. Phys.*, 10, 3953–3964, 2010.

4 King, S. M., Rosenoern, T., Shilling, J. E., Chen, Q., and Martin, S. T.: Increased cloud
5 activation potential of secondary organic aerosol for atmospheric mass loadings, *Atmos. Chem.*
6 *Phys.*, 9, 2959–2972, 2009.

7 Kreidenweis, S. M., Petters, M. D., and DeMott, P. J.: Deliquescence-controlled activation of
8 organic aerosols, *Geophys. Res. Lett.*, 33, L06801, doi: 10.1029/2005GL024863, 2006.

9 Kroll, J. H., Donahue, N. M., Jimenez, J. L., Kessler, S. H., Canagaratna, M. R., Wilson, K.
10 R., Altieri, K. E., Mazzoleni, L. R., Wozniak, A. S., Bluhm, H., Mysak, E. R., Smith, J. D.,
11 Kolb, C. E., Worsnop, D. R. Carbon oxidation state as a metric for describing the chemistry of
12 atmospheric organic aerosol, *Nature Chem.*, 3, 133-139, 2011.

13 Kuwata, M., Shao, W., Lebouteiller, R., Martin, S. T.: Classifying organic materials by oxygen-
14 to-carbon elemental ratio to predict the activation regime of cloud condensation nuclei (CCN),
15 *Atmos. Chem. Phys.*, 13, 5309-5324, 2013.

16 Massoli, P., Lambe, A. T., Ahern, A. T., Williams, L. R., Ehn, M., Mikkil a, J., Canagaratna,
17 M. R., Brune, W. H., Onasch, T. B., Jayne, J. T., Petäjä, T., Kulmala, M., Laaksonen, A., Kolb,
18 C. E., Davidovits, P., and Worsnop, D. R.: Relationship between aerosol oxidation level and
19 hygroscopic properties of laboratory generated secondary organic aerosol (SOA) particles,
20 *Geophys. Res. Lett.*, 37, L24801, doi: 10.1029/ 2010GL045258, 2010.

21 McFiggans, G., Artaxo, P., Baltensperger, U., Coe, H., Facchini, M., Feingold, G., Fuzzi, S.,
22 Gysel, M., Laaksonen, A., Lohmann, U., Mentel, T., Murphy, D., O’Dowd, C., Snider, J., and
23 Wein- gartner, E.: The effect of physical and chemical aerosol properties on warm cloud droplet
24 activation, *Atmos. Chem. Phys.*, 6, 2593–2649, 2006.

25 Petters, M. D., and Kreidenweis, S. M.: A single parameter representation of hygroscopicity
26 growth and cloud condensation nucleus activity, *Atmos. Chem. Phys.*, 7, 1961– 1971, 2007.

27 Petters, M. D., and Kreidenweis, S. M.: A single parameter representation of hygroscopicity
28 growth and cloud condensation nucleus activity – Part 2: Including solubility, *Atmos. Chem.*
29 *Phys.*, 8, 6273–6279, 2008.

- 1 Petters, M. D., and Kreidenweis, S. M.: A single parameter representation of hygroscopic
2 growth and cloud condensation nucleus activity – Part 3: Including surfactant partitioning,
3 *Atmos. Chem. Phys. Discuss.*, 12, 22687–22712, 2012.
- 4 Petters, M. D., H. Wex, E. Hallberger, L. Poulain, C. M. Carrico, G. McMeeking, S. M.
5 Kreidenweis, F. Stratmann, and Massling, A.: Towards closing the gap between hygroscopic
6 growth and activation for secondary organic aerosol – Part 2: Theoretical approaches, *Atmos.*
7 *Chem. Phys.*, 9, 3999–4009, 2009a.
- 8 Petters, M. D., Kreidenweis, S. M., Prenni, A.J., Sullivan, R. C., Carrico, C. M., Koehler, K.
9 A., and Ziemann, P.J.: Role of molecular size in cloud droplet activation, *J. Geophys. Res.*, 114,
10 D07209, doi: 10.1029/2008JD011532, 2009b.
- 11 Petters, M. D., Carrico, C. M., Kreidenweis, S. M., Prenni, A. J., DeMott, P.J., Collett Jr, J. L.,
12 and Hans Moosmüller, H.: Cloud condensation nucleation activity of biomass burning aerosol,
13 *J. Geophys. Res.*, 114, D22205, doi: 10.1029/2009JD012353, 2009c.
- 14 Poulain, L., Wu, Z., Petters, M. D., Wex, H., Hallbauer, E., Wehner, B., Massling, a.,
15 Kreidenweis, S. M. and Stratmann, F.: Towards closing the gap between hygroscopic growth
16 and CCN activation for secondary organic aerosols – Part 3: Influence of the chemical
17 composition on the hygroscopic properties and volatile fractions of aerosols, *Atmos. Chem.*
18 *Phys.*, 10(8), 3775–3785, doi:10.5194/acp-10-3775-2010, 2010.
- 19 Prausnitz, J. M., Lichtenhaler, R. M., Gomez de Azevedo, E.: **Molecular thermodynamics of**
20 **fluid-phase equilibria**, Prentice Hall; 3 edition, 1998.
- 21 Prenni, A. J., M. D. Petters, S. M. Kreidenweis, P. J. DeMott, and Ziemann, P.J.: Cloud droplet
22 activation of secondary organic aerosol, *J. Geophys. Res.*, 112, D10223, doi:
23 10.1029/2006JD007963, 2007.
- 24 Pruppacher, H.R., and J.D. Klett, *Microphysics of Clouds and Precipitation*, 954 pp., Kluwer
25 Acad., Norwell, Mass., 1997.
- 26 Psichoudaki, M. and Pandis, S. N.: Atmospheric aerosol water-soluble organic carbon
27 measurement: a theoretical analysis, *Environ. Sci. Technol.*, 47, 9791–8,
28 doi:10.1021/es402270y, 2013.
- 29 Rastak, N., Silvergren, S., Zieger, P., Wideqvist, U., Ström, J., Svenningsson, B., Maturilli, M.,
30 Tesche, M., Ekman, a. M. L., Tunved, P. and Riipinen, I.: Seasonal variation of aerosol water

1 uptake and its impact on the direct radiative effect at Ny-Ålesund, Svalbard, *Atmos. Chem.*
2 *Phys.*, 14, 7445–7460, doi:10.5194/acp-14-7445-2014, 2014.

3 Raymond, T. M. and Pandis, S. N.: Cloud activation of single-component organic aerosol
4 particles, *J. Geophys. Res.*, 107, 4787, doi:10.1029/2002JD002159, 2002.

5 Raymond, T. M. and Pandis, S. N.: Formation of cloud droplets by multicomponent organic
6 particles, *J. Geophys. Res.*, 108, 4469, doi:10.1029/2003JD003503, 2003.

7 Seinfeld, J. H. and Pandis, S. N.: *Atmospheric chemistry and physics: From air pollution to*
8 *climate change*, John Wiley, New York, 2006.

9 Shiraiwa, M., Berkemeier, T., Schilling-Fahnestock, K. A., Seinfeld, J. H., and Pöschl, U.:
10 *Molecular corridors and kinetic regimes in the multiphase chemical evolution of secondary*
11 *organic aerosol*, *Atmos. Chem. Phys.*, 14, 8323-8341, doi:10.5194/acp-14-8323-2014, 2014.

12 Shulman, M. L., Jacobson, M. C., Charlson, R. J., Synovec, R. E., and Young, T. E.: Dissolution
13 behavior and surface tension effects of organic compounds in nucleating cloud droplets. *J.*
14 *Geophys. Res.*, 23, 277-280, 1996.

15 Sotiropoulou, R. E. P., Medina, J., and Nenes, A.: CCN predictions: Is theory sufficient for
16 assessments of the indirect effect?, *Geophys. Res. Lett.*, 33, L05816,
17 doi:10.1029/2005GL025148, 2006.

18 Spracklen, D. V., Jimenez, J. L., Carslaw, K. S., Worsnop, D. R., Evans, M. J., Mann, G. W.,
19 Zhang, Q., Canagaratna, M. R., Allan, J., Coe, H., McFiggans, G., Rap, A., and Forster, P.:
20 Aerosol mass spectrometer constraint on the global secondary organic aerosol budget, *Atmos.*
21 *Chem. Phys.*, 11, 12109-12136, doi:10.5194/acp-11-12109-2011, 2011.

22 Suda, S. R., Petters, M. D., Matsunaga, A., Sullivan, R. C., Ziemann, P. J., and Kreidenweis,
23 S. M.: Hygroscopicity frequency distributions of secondary organic aerosols, *J. Geophys. Res.*
24 VOL. 117, D04207, doi:10.1029/2011JD016823, 2012.

25 Suda, S. R., Petters, M. D., Yeh, K., Strollo, C., Matsunaga, A., Faulhaber, A., Ziemann, P. J.,
26 Prenni, A. J., Carrico, C. M., Sullivan, R. C. and Kreidenweis, S. M.: Influence of functional
27 groups on organic aerosol cloud condensation nucleus activity, 2014.

28 Swietlicki, E., Hansson, H.-C., Hämeri, K., Svenningsson, B., Massling, A., Mcfiggans, G.,
29 McMurry, P. H., Petäjä, T., Tunved, P., Gysel, M., Topping, D., Weingartner, E., Baltensperger,
30 U., Rissler, J., Wiedensohler, A. and Kulmala, M.: Hygroscopic properties of submicrometer

1 atmospheric aerosol particles measured with H-TDMA instruments in various environments a
2 review, *Tellus B*, doi:10.1111/j.1600-0889.2008.00350.x, 2008.

3 Topping, D. O., McFiggans, G. B.: Tight coupling of particle size, number and composition in
4 atmospheric cloud droplet activation, *Atmos. Chem. Phys.*, 12, 3253–3260, 2012.

5 Topping, D.O., Barley, M., and McFiggans, G.: Including phase separation in a unified model
6 to calculate partitioning of vapours to mixed inorganic–organic aerosol particles, *Faraday*
7 *Discuss.*, 165, 273, 2013.

8 Twomey, S.: Pollution and the planetary albedo, *Atmos. Environ.*, 8, 1251–1256, 1974.

9 VanReken, T. M., Ng, N. L., Flagan, R. C., and Seinfeld, J. H.: Cloud condensation nucleus
10 activation properties of biogenic secondary organic aerosol, *J. Geophys. Res.*, 110, D07206,
11 doi:10.1029/2004JD005465, 2005.

12 Varutbangkul, V., Brechtel, F. J., Bahreini, R., Ng, N. L., Keywood, M. D., Kroll, J. H., Flagan,
13 R. C., Seinfeld, J. H., Lee, A., and Goldstein, A. H.: Hygroscopicity of secondary organic
14 aerosols formed by oxidation of cycloalkenes, monoterpenes, sesquiterpenes, and related
15 compounds, *Atmos. Chem. Phys.*, 6, 2367–2388, 2006.

16 Volkamer, R., Jimenez, J. L., San Martini, F., Dzepina, K., Zhang, Q., Salcedo, D., Molina, L.
17 T., Worsnop, D. R. and Molina, M. J.: Secondary organic aerosol formation from anthropogenic
18 air pollution: Rapid and higher than expected, *Geophys. Res. Lett.*, 33, L17811,
19 doi:10.1029/2006GL026899, 2006.

20 Wania, F., Lei, Y. D., Wang, C., Abbatt, J. P. D., and Goss, K.-U.: Novel methods for predicting
21 gas-particle partitioning during the formation of secondary organic aerosol, *Atmos. Chem.*
22 *Phys.* 14, 13189, 2014.

23 Wex, H., Petters, M. D., Carrico, C. M., Hallbauer, E., Massling, A., McMeeking, G. R.,
24 Poulain, L., Wu, Z., Kreidenweis, S. M., and Stratmann, F.: Towards closing the gap between
25 hygroscopic growth and activation for secondary organic aerosol: Part I – evidence from
26 measurements, *Atmos. Chem. Phys.*, 9, 3987–3997, 2009.

27 Zieger, P., Weingartner, E., Henzing, J., Moerman, M., de Leeuw, G., Mikkilä, J., Ehn, M.,
28 Petäjä, T., Clémer, K., van Roozendaal, M., Yilmaz, S., Frieß, U., Irie, H., Wagner, T.,
29 Shaiganfar, R., Beirle, S., Apituley, a., Wilson, K. and Baltensperger, U.: Comparison of
30 ambient aerosol extinction coefficients obtained from in-situ, MAX-DOAS and LIDAR

1 measurements at Cabauw, Atmos. Chem. Phys., 11, 2603–2624, doi:10.5194/acp-11-2603-
2 2011, 2011.
3
4
5
6
7
8
9
10

1

Table 1. Simplified descriptions of organic mixture solubilities.

Mixture model	Number of components	Solubility Presentation	Other input parameters
Complete dissolution	1	$c_{sat} \rightarrow \infty$ ¹	M_{org}, ρ_{org}
κ	1	$\kappa(c_{sat}, M_{org}, \rho_{org})$	-
Soluble fraction ε_{eff}	2	$c_{sat,1} \rightarrow \infty$ $c_{sat,2} = 0$	M_{org}, ρ_{org}

2

¹ c_{sat} the solubility (saturation concentration) in aqueous solution.

1 **Table 2.** Solubility distributions of the organic mixtures considered in this study.

Distribution ¹	Shape	Number of components	$[c_{sat,min}, c_{sat,max}]$ ² (g L ⁻¹)
1	Flat, log <i>c</i> -axis		
2	Flat, linear <i>c</i> -axis		Low: [10 ⁻⁵ , 10 ³]
3	Log. increasing	3, 5, 10, 100	Mid: [0.1, 10 ³]
4	Linear increasing		High: [10, 10 ³]
5	Log. decreasing		
6	Linear decreasing		

18 ¹For all solubility distributions two assumptions about the organic phase activity coefficients:
19 (1) ideal mixture and (2) unity activity (see text for details).

20 ² c_{sat} the solubility (pure component saturation concentration) in aqueous solution.

1 **Table 3.** Properties of water and organic compounds used in Köhler curve calculations (see Eq.
2 1).

Property (unit) ¹	Water	Organic <i>i</i>
ρ (kg m ⁻³)	1000	1500
σ (N m ⁻¹)	0.073	-
M (kg mol ⁻¹)	0.018	0.18

3 ¹ These properties were chosen based on literature on the effective molar masses and densities
4 determined for laboratory SOA (Engelhart et al., 2008; Asa-Awuku et al., 2010), and assumed
5 to be same for every organic compound *i*.

6

7

8

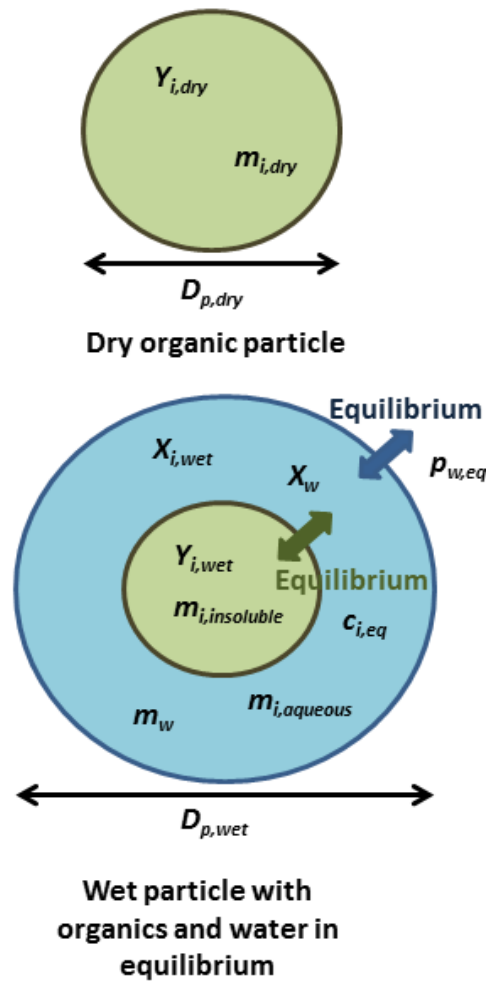
9

10

11

1

2



3

4

5 **Figure 1.** Schematic of the conceptual model used in the equilibrium composition calculations.

6 The dry particle is assumed to consist of n organic compounds, each denoted with a subscript

7 i . The wet particle is assumed to consist of a dry organic (insoluble) phase and an aqueous phase

8 with water and dissolved organics. The aqueous phase is assumed to be in equilibrium with the

9 ambient water vapour. Y refers to mole fractions in the organic phase, X to mole fractions in the

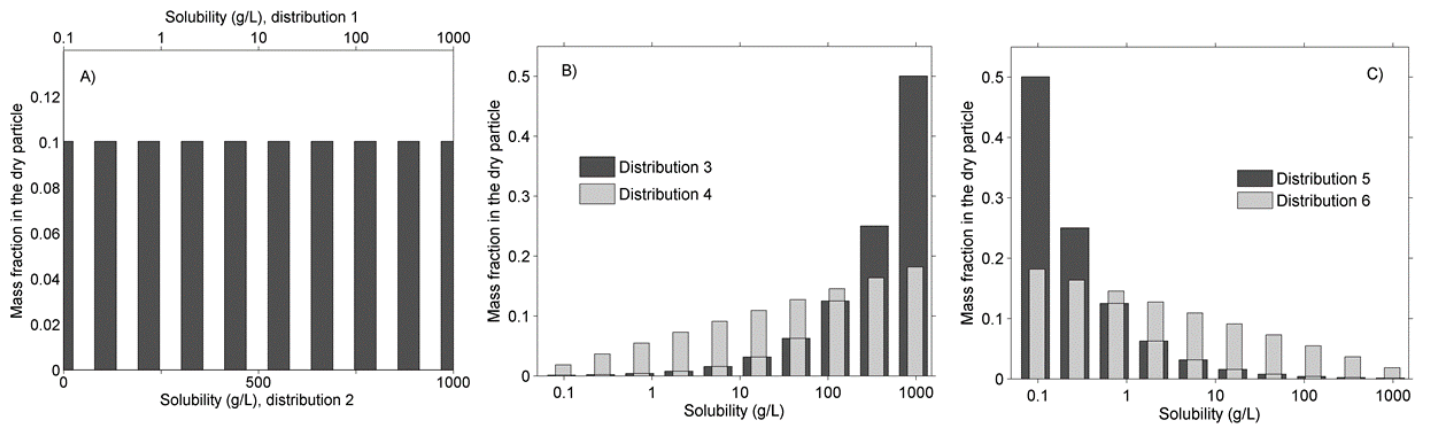
10 aqueous phase, and m to the masses of the organic constituents and water. $c_{i,eq}$ refers to the

11 equilibrium concentration of each organic compound in the aqueous solution and $p_{w,eq}$ to the

12 equilibrium vapour pressure of water above the aqueous solution.

13

14

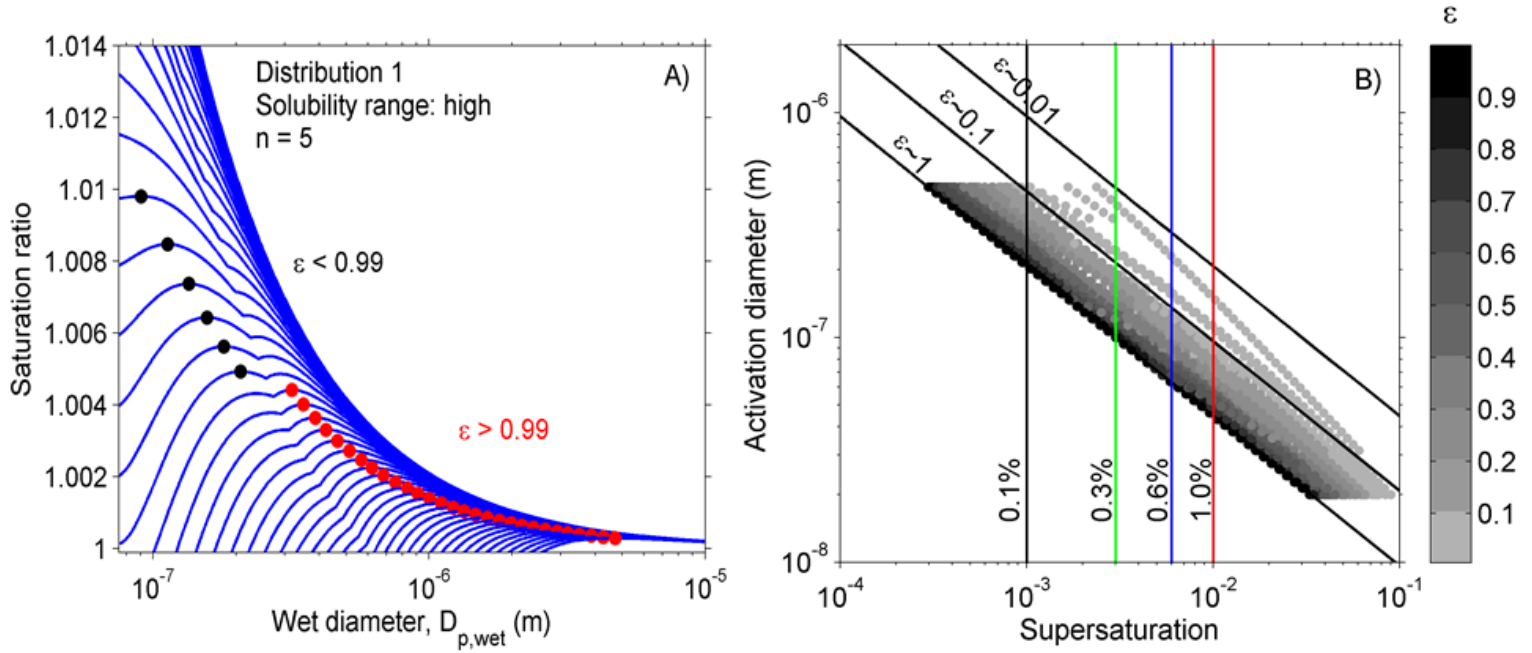


1

2 **Figure 2.** Examples of solubility distributions used in the calculations for saturation
 3 concentrations ranging from 0.1 to 1000 g L⁻¹. (a) Linear and logarithmic flat distributions; (b)
 4 Linear and logarithmic increasing distributions; (c) Linear and logarithmic decreasing
 5 distributions. The numbers of the distributions refer to the numbering in Table 2 (see Sect.
 6 2.1.4).

7

1

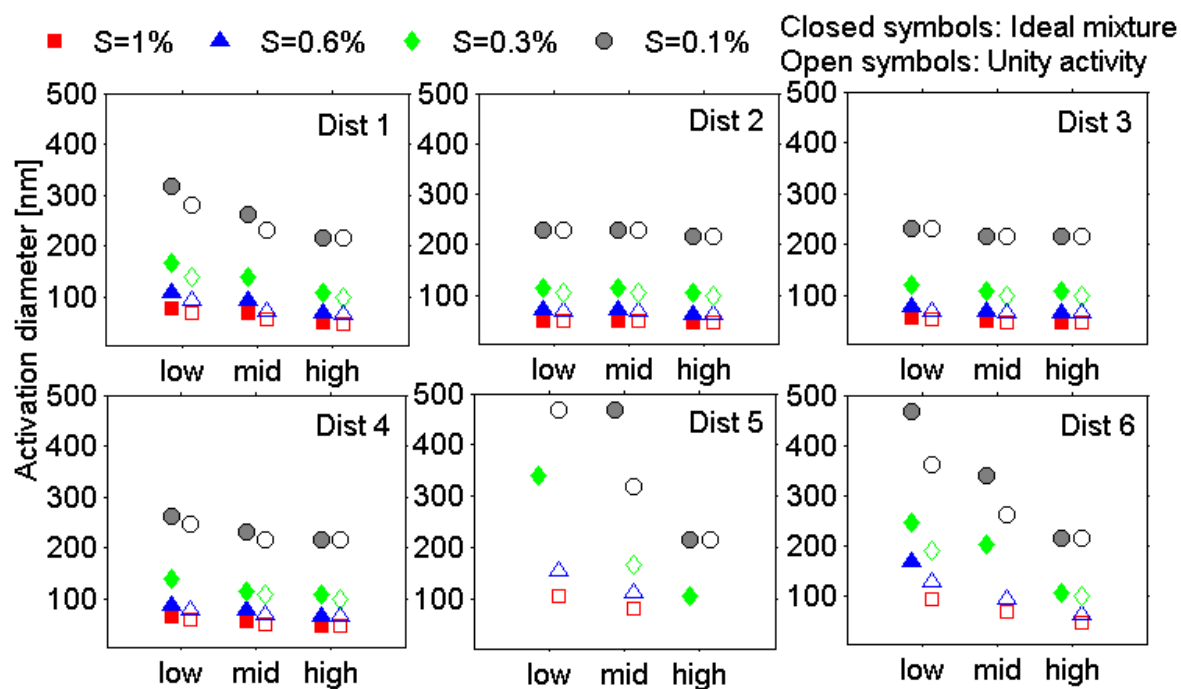


2

3 **Figure 3.** a) Examples of Köhler curves for the flat logarithmically spaced solubility
 4 distribution (Distribution 1 in Table 2) with $n = 5$ and the high solubility range (Table 2). The
 5 dots indicate the point of activation, black indicating incomplete dissolution ($\epsilon < 1$) and red
 6 complete dissolution ($\epsilon = 1$). b) The activation points determined from the model calculations
 7 (in total 7200 Köhler curves, see Table 2), corresponding to in total 5957 points in the activation
 8 dry diameter vs. supersaturation space. Also the dependence of the dissolved fraction at the
 9 point of activation is illustrated.

10

1

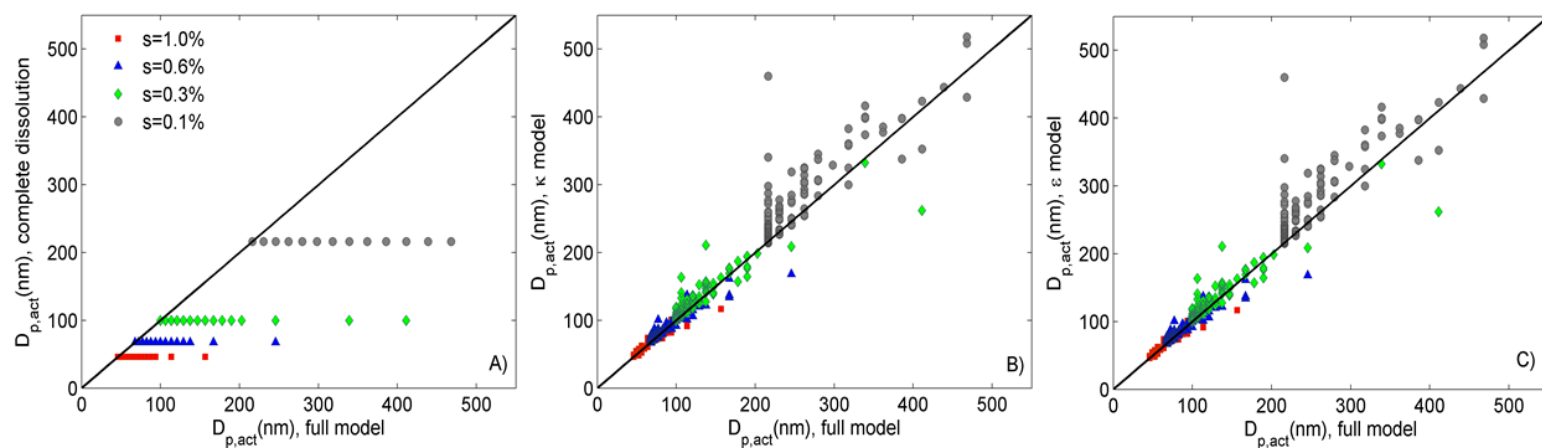


2

3 **Figure 4.** The dependence of the activation diameter for four different supersaturations ($s =$
4 1%, 0.6%, 0.3% and 0.1%, see also Fig. 3b) on the solubility range for the solubility
5 distributions outlined in Table 2 for $n = 5$, and the two assumptions about the organic phase
6 activity.

7

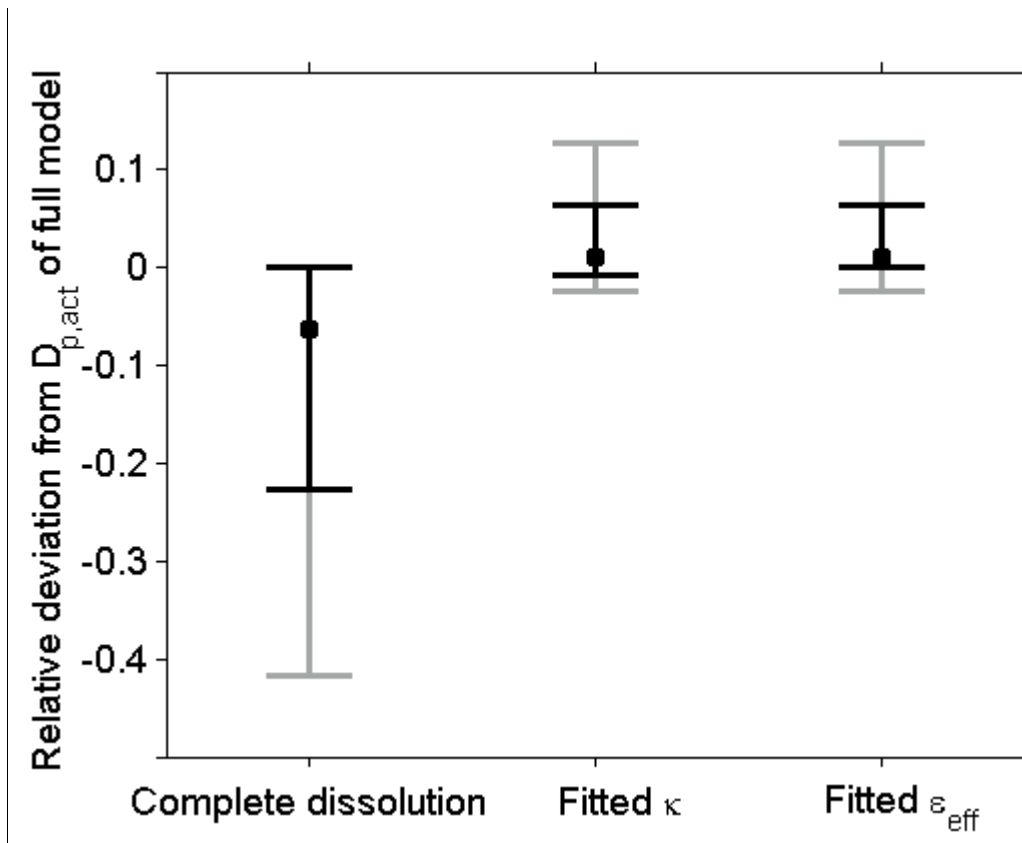
1



2

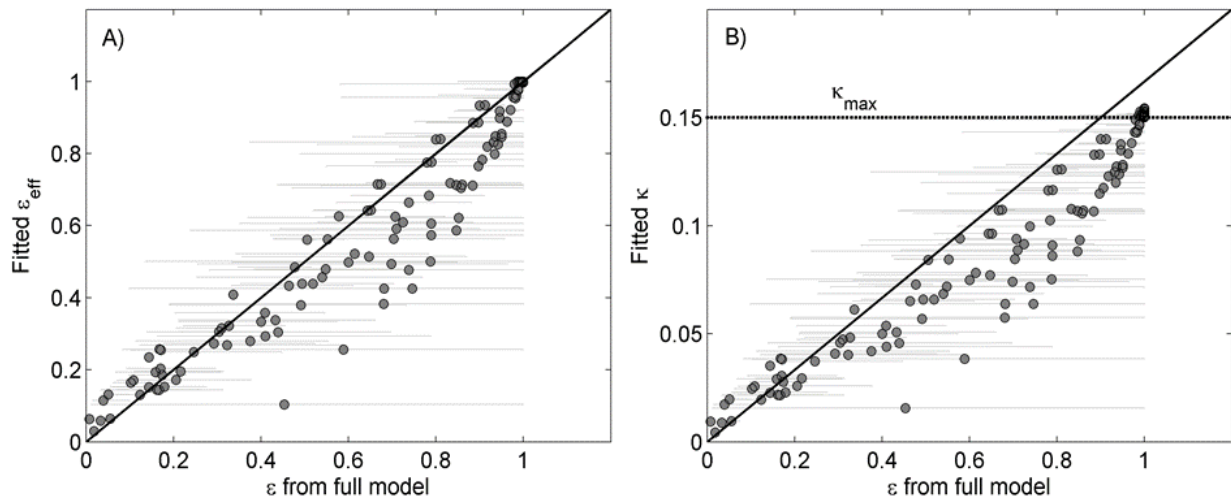
3 **Figure 5.** The activation diameter calculated using the solubility distributions (Table 2, referred
4 to as the full model) and the simplified dissolution descriptions (Table 1) a) complete
5 dissolution assumption; b) the hygroscopicity parameter κ ; c) the soluble fraction ϵ_{eff} to describe
6 the solubility of the organic mixture. The symbols correspond to the best fits to the full model
7 data. The black line shows the 1:1 correspondence between the two data sets.

8



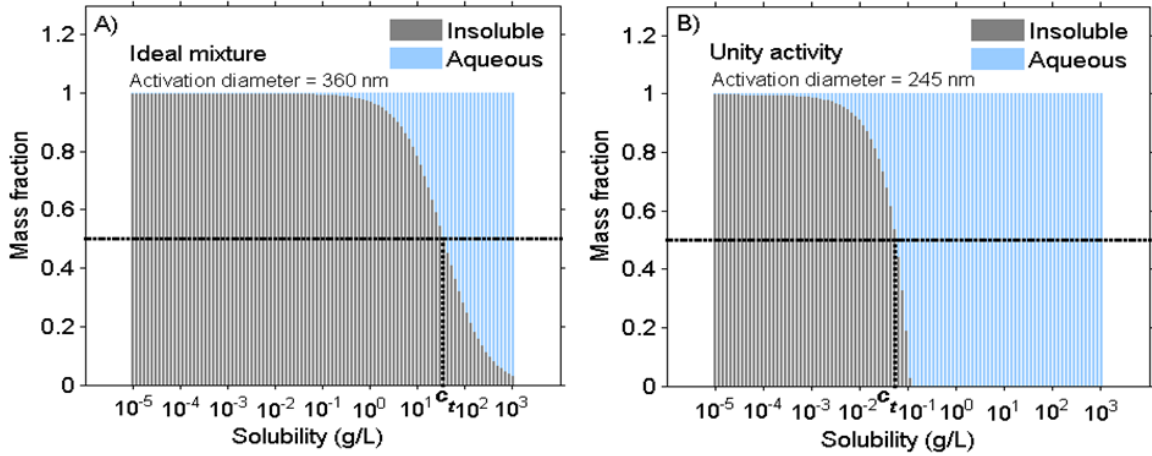
1
2
3
4
5
6
7

Figure 6. The performance of the simplified solubility representations (see Table 1) in predicting the activation diameter for a given supersaturation as compared with the full model. The black bars depict the 25- and 75-percentiles and the gray bars the 10- and 90-percentiles.



1
2
3
4
5
6
7
8
9

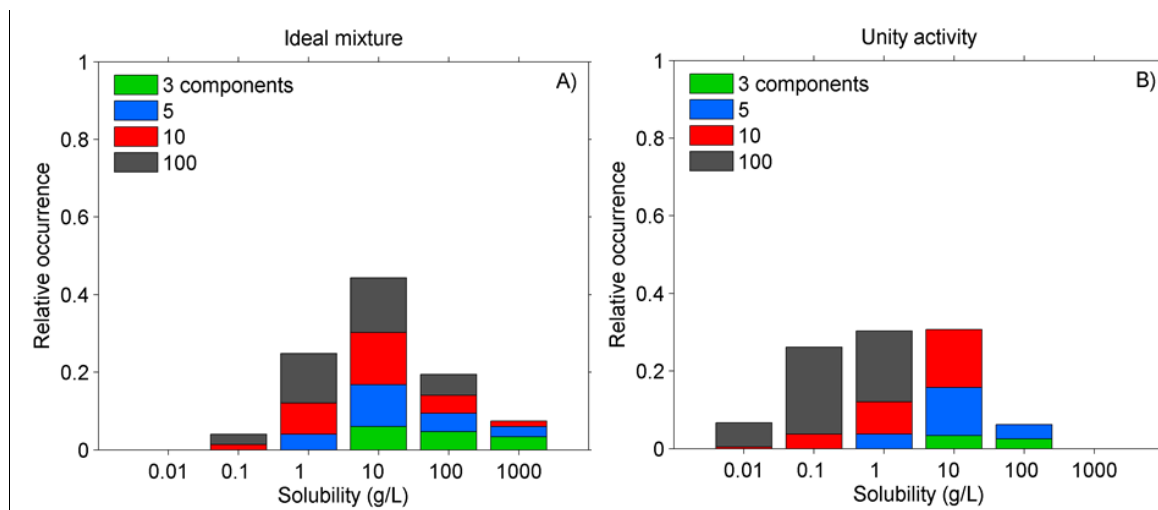
Figure 7. a) The fitted soluble fraction ε_{eff} as function of the true dissolved fraction ε for each considered mixture (see Table 2). Symbols: mean ε over all activation points. Grey lines: the range of ε values at the different activation points for a given mixture; b) The fitted κ values as a function as function of the true dissolved fraction ε for each considered mixture. The red dashed line denotes the limit of $\kappa_{max} = 0.15$ which applies to all the studied mixtures. 1:1 lines are also indicated.



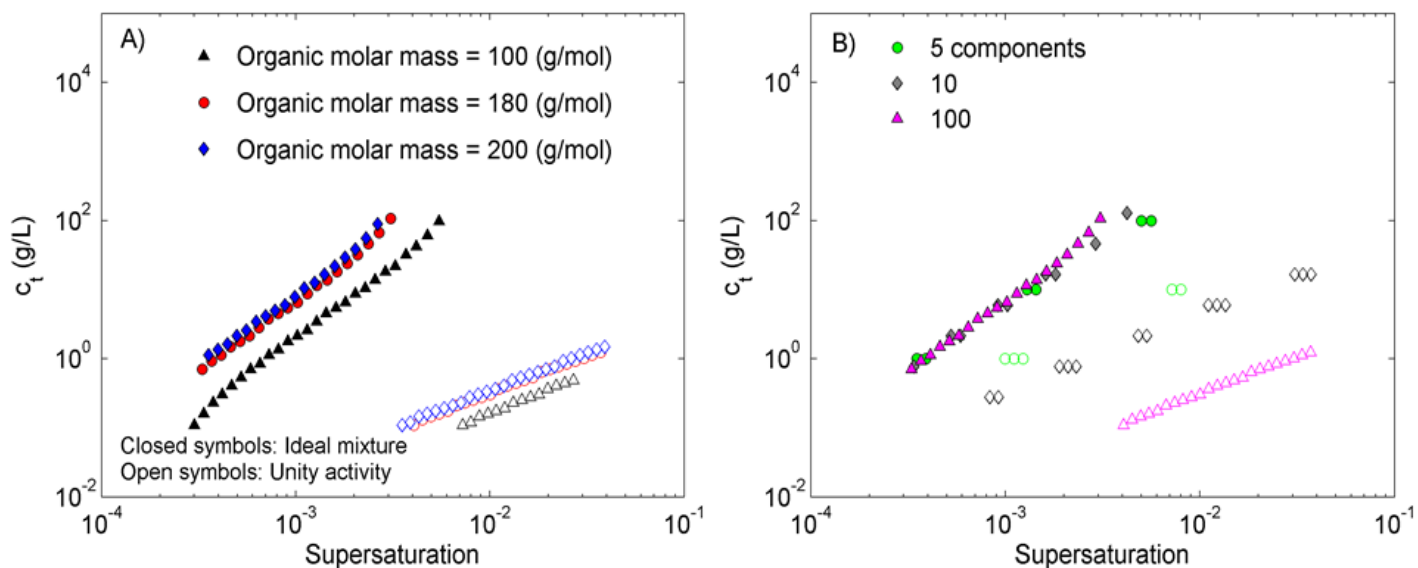
1

2 **Figure 8.** The dissolution behavior of the organic mixture corresponding to Distribution 1 with
 3 $n = 100$ and the low solubility range (see Table 2) at the activation point for $s_c = 0.1\%$ (see Fig.
 4 4). The figures depict the distribution of material in each solubility bin between the aqueous
 5 and the insoluble organic phases for the two different assumptions about the organic phase
 6 activity. c_i refers to the 50%-point of the partitioning (Eqs. 24-25). a) The ideal organic mixture;
 7 b) the unity activity assumption.

8

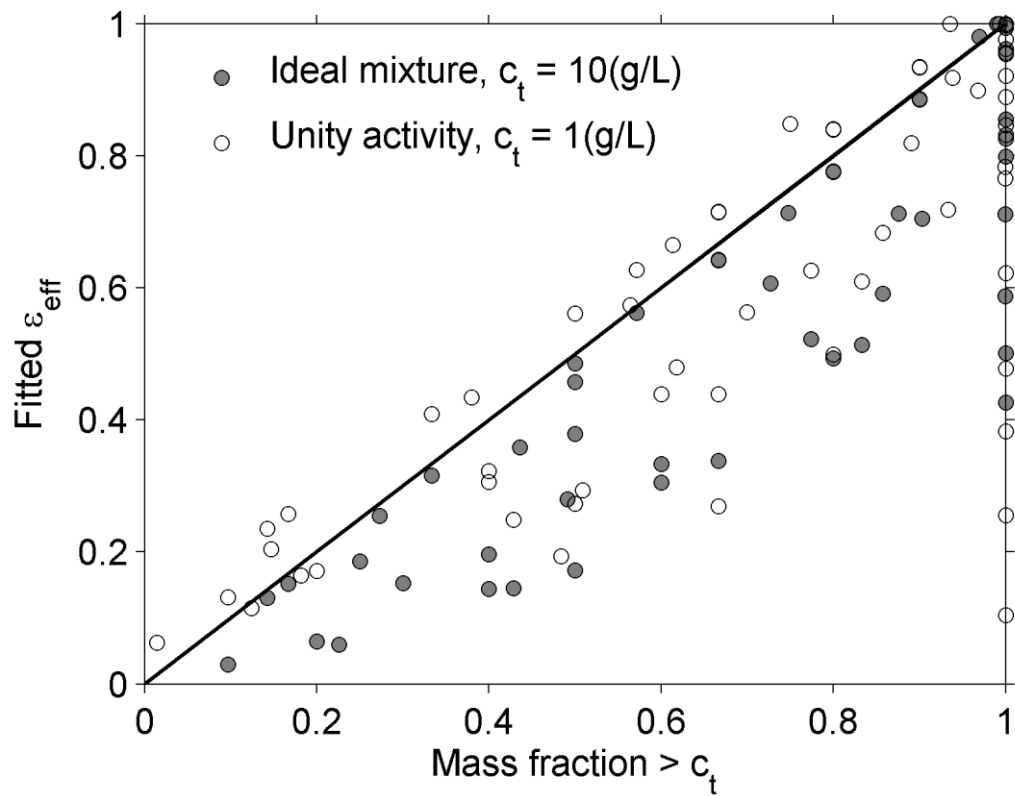


1
 2 **Figure 9.** The distributions of the c_i values (i.e. the 50% partitioning point, see Eqs. 24-25) at
 3 the point of activation for all the considered mixtures (Table 2) and activation points, and the
 4 two assumptions about the organic phase activity. a) The ideal organic mixture; b) the unity
 5 activity assumption.
 6



1
 2 **Figure 10.** Sensitivity of the c_t values to supersaturation, molar mass and number of
 3 components for Distribution 1 with the mid solubility range (see Table 2). Only points with
 4 limited solubility ($0 < \varepsilon < 1$) at the point of activation are included.

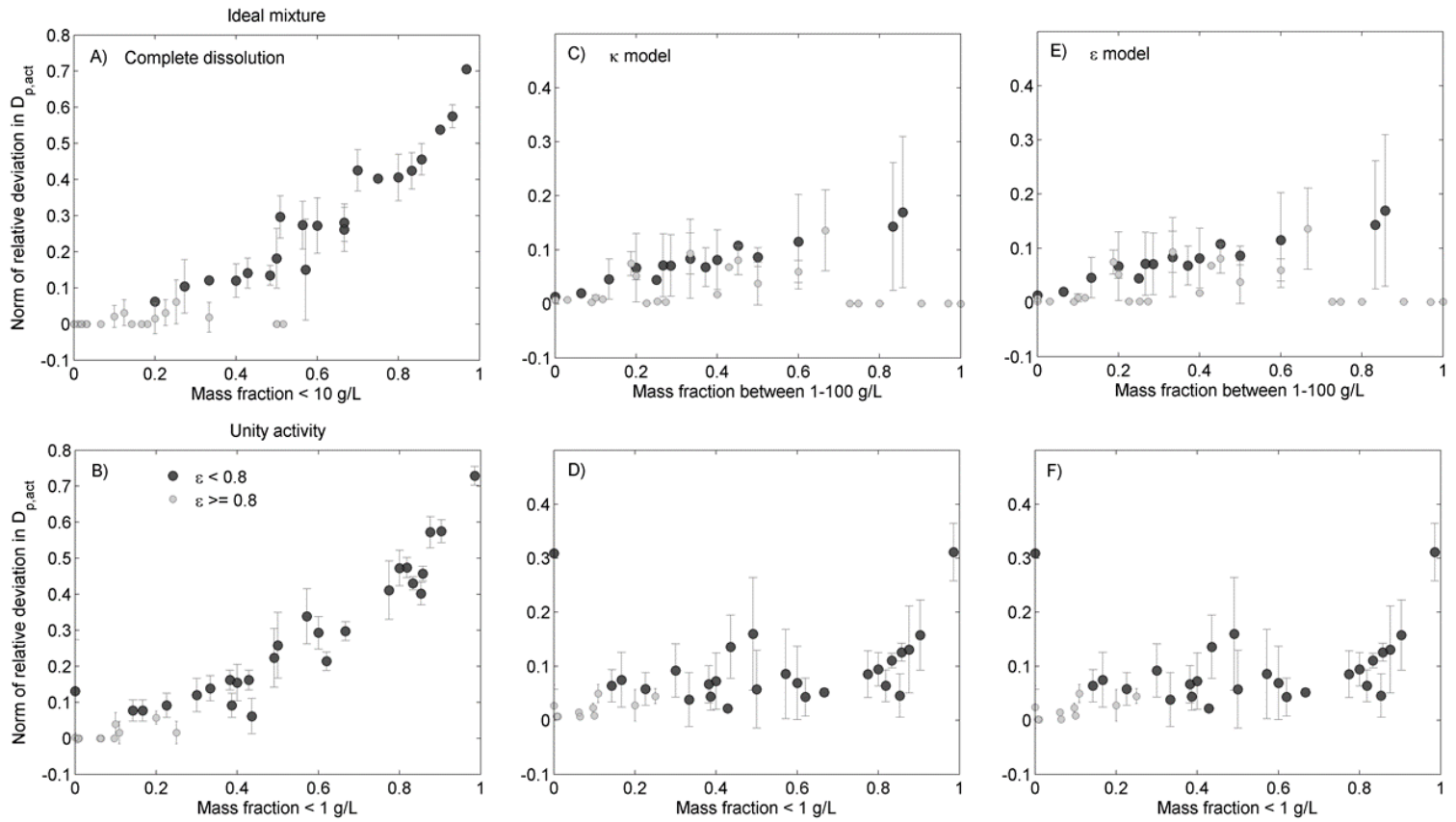
5



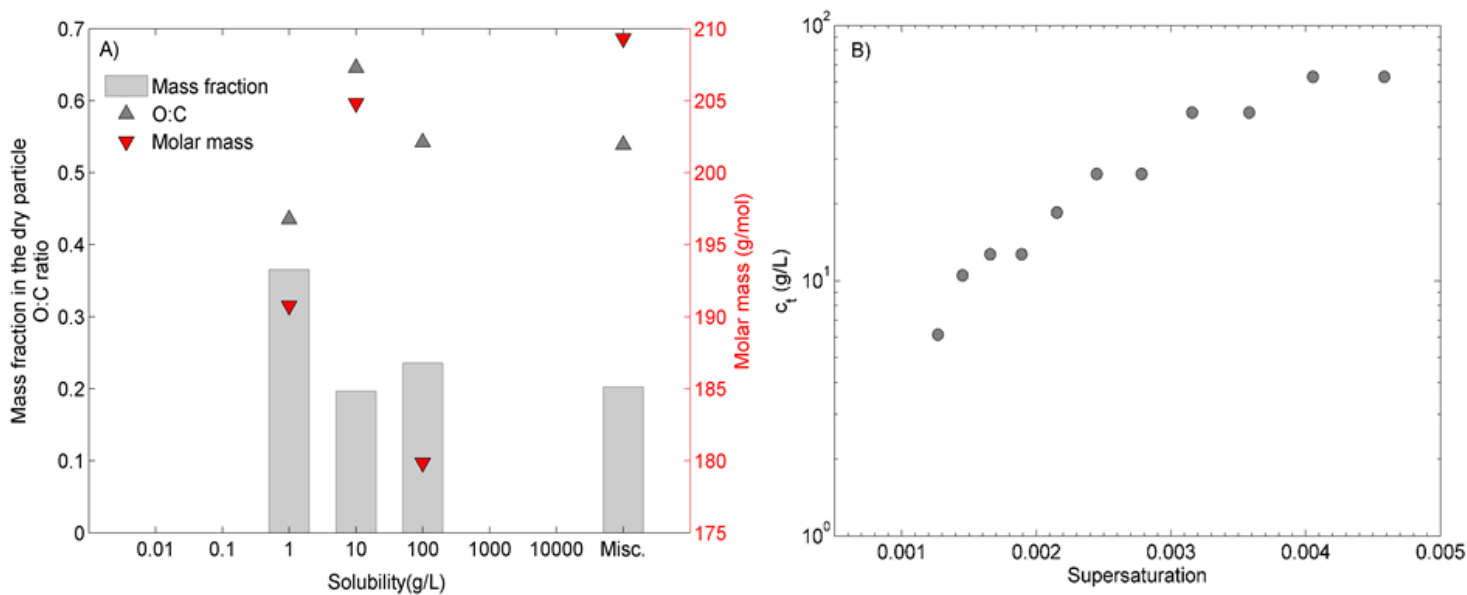
1

2 **Figure 11.** The relationship between the fitted dissolved fraction ϵ_{eff} at the point of activation
 3 and the mass fraction over the median c_t for all the considered mixtures and $s = 1\%$, 0.6% , 0.3%
 4 and 0.1% . Closed symbols: the ideal organic mixture assumption (median $c_t = 10 \text{ g L}^{-1}$). Open
 5 symbols: the unity activity assumption for the organics (median $c_t = 1 \text{ g L}^{-1}$).

6



1 **Figure 12.** The solubility distribution properties best explaining the performance of the three
 2 simplified solubility models, illustrated with the norm of the relative deviation of $D_{p,act}$ as
 3 compared with the full model predictions for $s = 1\%$, 0.6% , 0.3% and 0.1% . The performance
 4 of the complete dissolution assumption as a function of the mass fraction with solubilities below
 5 the median c_t for a) the ideal organic mixture (median $c_t = 10 \text{ g L}^{-1}$) and b) the unity organic
 6 activity (median $c_t = 1 \text{ g L}^{-1}$) assumptions. The performance of the κ model as a function of the
 7 mass fraction with solubilities c) between $1\text{-}100 \text{ g L}^{-1}$ for the ideal organic mixture assumption;
 8 d) below 1 g L^{-1} for the unity organic activity assumption. The performance of the ε model as a
 9 function of the mass fraction with solubilities e) between $1\text{-}100 \text{ g L}^{-1}$ for the ideal organic
 10 mixture assumption; f) below 1 g L^{-1} for the unity organic activity assumption. The points close
 11 to to complete dissolution ($\varepsilon \geq 0.8$) are shown with lighter grey than the rest of the points. The
 12 error bars represent the variability with supersaturation and particle size.
 13



1
 2 **Figure 13.** A) An example of a solubility distribution for SOA generated in dark-ozonolysis of
 3 α -pinene SOA. The expected composition has been taken from Chen et al. (2008), and the pure-
 4 component solubilities have been calculated with the SPARC prediction system (e.g. Wania et
 5 al., 2014 and references therein). “Misc.” refers to completely miscible components. B) The
 6 dependence of the c_t value at the point of activation on supersaturation assuming an ideal
 7 organic mixture, with only points corresponding to limited solubility ($0 < \varepsilon < 1$) displayed.