Prediction of cloud condensation nuclei activity for organic compounds using functional group contribution methods

Markus D. Petters[1*], Sonia M. Kreidenweis[2], Paul J. Ziemann[3]

[1]{Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC, USA.}

[2]{Department of Atmospheric Sciences, Colorado State University, Fort Collins, CO, USA.}

[3]{Department of Chemistry and Biochemistry, Colorado University, Boulder, CO, USA.}

Correspondence to: Markus D. Petters (markus_petters@ncsu.edu)

Revised version submitted to: *Geoscientific Model Development*

Date: Dec 2, 2015

Keywords

Cloud condensation nuclei, organic aerosol, thermodynamic modeling, UNIFAC, functional group composition

Abstract

A wealth of recent laboratory and field experiments demonstrate that organic aerosol composition evolves with time in the atmosphere, leading to changes in the influence of the organic fraction to cloud condensation nuclei (CCN) spectra. There is a need for tools that can realistically represent the evolution of CCN activity to better predict indirect effects of organic aerosol on clouds and climate. This work describes a model to predict the CCN activity of organic compounds from functional group composition. Following previous methods in the literature, we test the ability of semi-empirical group contribution methods in Köhler theory to predict the effective hygroscopicity parameter, kappa. However, in our approach we also account for liquid-liquid phase boundaries to simulate phase-limited activation behaviour. Model evaluation against a selected database of published laboratory measurements demonstrates that kappa can be predicted within a factor of two. Simulation of homologous series is used to identify the relative effectiveness of different functional groups in increasing the CCN activity of weakly functionalized organic compounds. Hydroxyl, carboxyl, aldehyde, hydroperoxide, carbonyl, and ether moieties promote CCN activity while methylene and nitrate moieties inhibit CCN activity. The model can be incorporated into scale-bridging testbeds such as the Generator of Explicit Chemistry and Kinetics of Organics in the Atmosphere to evaluate the evolution of kappa for a complex mix of organic compounds and to develop suitable parameterizations of CCN evolution for larger scale models.

Introduction

Organic compounds are an important contributor to the atmospheric submicron aerosol (Jimenez et al., 2009). The organic fraction is projected to increase in the future due to the confluence of a decreasing sulfate and nitrate burden and increases in the global secondary organic aerosol burden (Heald et al., 2008). An important unanswered question is how the organic fraction influences the aerosol's ability to serve as cloud condensation nuclei (CCN), and in turn modulates climate via indirect effects of aerosols on clouds and precipitation (Andreae and Rosenfeld, 2008). Realistic prescribed variations in secondary organic aerosol hygroscopicity have demonstrable impacts on CCN number concentration (Mei et al., 2013) and can change the simulated global aerosol indirect forcing (AIF) by ~1/6 of the AIF simulated in a control case (Liu and Wang 2010). To obtain a prognostic understanding of the contribution of the organic fraction to indirect aerosol forcing in future climates, models need improved schemes that map simulated organic aerosol composition to hygroscopicity and CCN activity.

Several organic aerosol types (e.g. freshly emitted diesel oil particles or first generation oxidation products of sesquiterpenes) consist of mostly hydrophobic hydrocarbon chains with few functional groups attached. Pure hydrocarbons with carbon number less than C_{30} are expected to be semi-volatile and in the liquid phase. Over time the compounds evolve by functionalization, fragmentation and oligomerization (Kroll and Seinfeld, 2008, Ziemann and Atkinson, 2012). As functional groups are added to the carbon chain the products usually, but not always, become less volatile (Goldstein et al., 2007), more dense (Kuwata et al., 2012), more viscous (Sastri and Rao, 1992), and more CCN active (Suda et al., 2014).

Laboratory (George and Abbatt, 2010, Poulain et al., 2010, Cappa et al., 2011, Massoli et al., 2010, Lambe et al., 2011, Duplissy et al., 2011, Kuwata et al., 2013, Rickards et al., 2013, Suda et al., 2014) and field studies (Jimenez et al., 2008, Chang et al., 2008, Mei et al., 2013) have demonstrated a robust link between the aerosol oxidation state and the ability of the organic fraction to promote hygroscopic water uptake and CCN activity. Proxies from mass spectrometry such as the fragmentation peak f_{44} or the atomic oxygen-to-carbon ratio are often used to model the increase in hygroscopicity. However, these correlations exhibit significant variability

between studies and break down when applied at the compound level (Rickards et al., 2013, Suda et al., 2014).

Chemistry models are already capable of simulating the molecular identities of species present in the condensed phase during multi-day evolution of diluting air-parcels (Lee-Taylor et al., 2015). Mapping this speciated aerosol composition to the aerosol hygroscopicity should ultimately permit quantification of changes in CCN number concentration (provided that the size distribution is also simulated) and associated effects on clouds and climate. Thermodynamic models should be able to predict CCN activity. Many thermodynamic models have made use of activity coefficients predicted by the universal functional group activity coefficient (UNIFAC) group contribution method (Fredenslund et al., 1975). Several investigators have compared UNIFAC predictions of organic aerosol water content to experimental data (Saxena and Hildemann, 1997, Ming and Russell, 2001, Peng et al., 2001, Choi and Chan, 2002, Mochida and Kawamura, 2004, Marcolli and Peter, 2005, Moore and Raymond, 2008). Some of these comparisons prompted proposed revisions of specific group interaction parameters, e.g. [OH] and [H_2O]. Several thermodynamic models that treat complex phase equilibria of multifunctional multicomponent organic mixtures are based on UNIFAC activity coefficients (Ming and Russell, 2002, Raatikainen and Laaksonen, 2005, Topping et al., 2005, Amundson et al., 2007, Zuend et al., 2008, Compernolle et al., 2009). The development of these models has been driven by the need to enable predictions over a wide range of conditions and compositions, including the effect of liquid-liquid phase separation on gas-to-particle partitioning (Zuend and Seinfeld, 2012, Toping et al., 2013). The prediction of CCN activity of organic compounds has received less attention. Rissman et al. (2007) used the aerosol diameter dependent equilibrium model (ADDEM, Topping et al., 2005) with an underlying UNIFAC core to predict the relationship between critical supersaturation and dry for several dicarboxylic acid aerosols. To our knowledge no study to date has systematically focused on the prediction of CCN activity from thermodynamic models.

Here we build on this body of work to predict the contribution of a compound with known chemical structure to the CCN activity of a particle of known size. The proposed model uses the UNIFAC equations (Fredenslund et al., 1975) with group interaction parameters form Hansen et

al., (1991), Raatikainen and Laaksonen (2005) and Compernolle et al. (2009) to model activity coefficients and free energy of mixing. Liquid-liquid phase boundaries are determined using the area method of Eubank et al. (1992). Molecular volume is estimated from elemental composition and adjustments for functional group composition using the approach of Girolami (1994). The relationship between critical supersaturation and dry diameter is then predicted using Köhler theory (Seinfeld and Pandis, 2006). The basic model mechanics are similar to those employed in multicomponent phase equilibrium models (Ming and Russell, 2002, Raatikainen and Laaksonen, 2005, Topping et al., 2005, Amundson et al., 2007, Zuend et al., 2008) but limited in scope to binary compositions and with focus on accurately representing phase and water activity at conditions relevant at the point of CCN activation only. These predictions are validated by manually mapping chemical composition to UNIFAC groupings and comparing modeled CCN activity against observations from a compiled library of recently published CCN data of mostly weakly oxidized hydrocarbons containing a mixture of alcohol, carbonyl, aldehyde, ether, carboxyl, nitrate, and hydroperoxide moieties. The model is used to predict how the addition of one or more functional groups to otherwise similar molecules promotes CCN activity. Envisioned application to multi-component aerosols and contrasts with more complete thermodynamic models are discussed.

Model Description

Köhler theory

The saturation ratio over a curved droplet is given by the Köhler equation

$$S = a_w \cdot \exp\left(\frac{4\sigma_{s/a}(T)M_w}{\rho_w RTD}\right), \quad (1)$$

where a_w is the water activity, $\sigma_{s/a}$ is the surface tension of the solution/air interface, T is temperature, M_w is the molecular weight of water, ρ_w is the density of pure water, R is the universal gas constant, and D is the wet drop diameter. Water activity depends on the water content and the amounts and identities of solutes in the nucleus. The principle water content variable used in this work is the mole fraction

$$x_w = \frac{n_w}{n_w + \sum_i n_{s,i}}, \quad (2)$$

where x_w is the mole fraction of water, n_w and $n_{s,i}$ are the number of moles of water and solutes, and i is the number of dry components. The wet drop diameter can be calculated from x_w if the dry diameter, D_d, is specified and it is assumed that the particle is spherical and that the volume of water and solute are additive:

$$D = \langle (x_w - 1)^{-1} (x_w - x_w \sum_i (\epsilon_i v_w v_{s,i}^{-1}) - 1) D_d^3 \rangle^{1/3}. \quad (3)$$

In Eq. (3) v_w and $v_{s,i}$ are the molar volume of the water and solutes and ε_i are the volume fractions in the dry particle. Eq. (3) is obtained by rearranging Eq. (7) in Petters et al. (2009a). The critical supersaturation required for an aqueous solution droplet to activate into a cloud droplet is found by combining Eqs. (1) and (3) and finding the x_w (or D) that maximizes s_c

$$s_c = \left\{ \max_{x_w \in [0,1]} \left[a_w \cdot \exp\left(\frac{4\sigma_{s/a}(T) M_w}{\rho_w RT \langle (x_w - 1)^{-1} (x_w - x_w \sum_i (\epsilon_i v_w v_{s,i}^{-1}) - 1) D_d^3 \rangle^{1/3}} \right) \right] \right\} \cdot 100\%, \quad (4)$$

where s_c is the critical supersaturation in %. The variables that control s_c are v_s, a_w, and $\sigma_{s/a}$. In this work it is assumed that surface tension is that of pure water. Discussion on this and other assumptions are provided at the end of this section. First the prediction of v_s and a_w for organic compounds with known chemical structure is described.

Molar Volume

Molar volume is calculated from the molecular formula using the method of Girolami (1994). Each element is assigned a relative volume based on its location in the periodic table. The elemental volumes are summed and scaled by a constant factor to compute v_s. If the oxygen is bound in the form of alcohol [OH] or carboxyl [C(=O)OH] moieties, the actual v_s is smaller due to intramolecular bonding. Therefore v_s is decreased by 10% for each [OH] or [C(=O)OH] group but by no more than 30% of the molar volume derived from the elemental composition. Girolami (1994) tested this method for 166 liquids and reports agreement with observations $v_s \sim \pm 10\%$. Barley et al. (2012) reviewed the performance of various methods for predicting molar volume using a test set of 56 multifunctional organic compounds and report similar scatter.

Water activity

Water activity is related to the mole fraction via

$$a_w = \gamma_w x_w, \quad (5)$$

where γ_w is the activity coefficient of water. Activity coefficients are estimated using the semi-empirical group contribution method UNIFAC (Fredenslund et al., 1975). The UNIFAC model describes a liquid solution that consists of i components. Each component is divided into k groups. The activity coefficient of component i in solution (γ_i) has contributions from combinatorial (γ^C) and residual parts (γ^R)

$$\ln\gamma_i = \ln\gamma_i^C + \ln\gamma_i^R. \quad (6)$$

The combinatorial part is computed via

$$\ln\gamma_i^C = \ln\frac{\Phi_i}{x_i} + \frac{z}{2} q_i \ln\frac{\theta_i}{\Phi_i} + l_i - \frac{\Phi_i}{x_i}\sum_j x_j l_j \quad (7a)$$

$$l_i = \frac{z}{2}(r_i - q_i) - (r_i - 1); z = 10 \quad (7b)$$

$$\theta_i = \frac{q_i x_i}{\sum_j q_j x_j}; \Phi_i = \frac{r_i x_i}{\sum_j r_j x_j} \quad (7c)$$

$$r_i = \sum_k v_k^{(i)} R_k \,; q_i = \sum_k v_k^{(i)} Q_k. \quad (7d)$$

In Eqs. (7), x_i is the mole fraction of component i, θ_i and Φ_i are the average surface and segment fraction, z is the lattice coordination number, $v^{(i)}_k$ is the number of groups of type k in component i, R_k and Q_k are the group volume and surface area parameters derived from Bondi (1964), and r_i and q_i are the normalized Van-der-Waals volume and surface area. The summation i or j is over all components in the mixture, including component i.

The residual part is computed via

$$\ln\gamma_i^R = \sum_k v_k^{(i)}\left[\ln\Gamma_k - \ln\Gamma_k^{(i)}\right] \quad (8a)$$

$$\ln\Gamma_k = Q_k - \left[1 - \ln\left(\sum_m \Theta_m \Psi_{mk}\right) - \sum_m \frac{\Theta_m \Psi_{km}}{\sum_n \Theta_n \Psi_{nm}}\right] \quad (8b)$$

$$\Theta_m = \frac{Q_m X_m}{\sum_n Q_n X_n} \quad (8c)$$

$$X_m = \frac{\sum_i v_m^{(i)} x_i}{\sum_i \sum_k v_k^{(i)} x_i} \quad (8d)$$

$$\Psi_{mn} = \exp\left(-\frac{a_{mn}}{T}\right) \quad (8e)$$

In Eqs (8), a_{mn} are empirically determined parameters, Ψ_{mn} is the group interaction parameter of group m with n, X_m is the mole fraction of group m in the mixture, Θ_m is the area fraction of group m, Γ_k is the group residual activity coefficient, and $\Gamma^{(i)}_k$ is the residual activity coefficient of group k in a reference solution containing only molecules of type i. Eqs. (8) are also used to compute $\Gamma^{(i)}_k$. The summation n or m is over all different groups in the mixture, and the summation k is over all groups in component i.

Groups within UNIFAC are represented as main groups and subgroups. The main groups evaluated in this work are alkane [CH_n], alcohol [OH], water [H_2O], carbonyl [$CH_nC(=O)$], aldehyde [HC(=O)], ether [$CH_n(O)$], carboxyl [C(=O)OH], nitrate [CH_nONO_2], and hydroperoxide [$CH_n(OOH)$]. Interaction parameters a_{mn} between the main groups that are used in this work are tabulated in Table S1. Some of the main groups have several subgroups, with each subgroup having unique volume and surface area parameters R_k and Q_k. These are summarized in Table S2.

Phase Equilibrium

For some x_w liquid-liquid phase separation can occur. The normalized Gibbs free energy of the mixture, defined as the actual Gibbs free energy divided by the thermal energy, is needed to compute the number of thermodynamically stable phases in the system. For a binary system consisting of water (w) and a single solute (s), Gibbs energy is calculated from the activity coefficients via standard thermodynamic relationships (Prausnitz et al., 1999, Petters et al., 2009)

$$\Delta g^{mix} = \Delta g^{ideal} + \Delta g^{excess} \quad (9a)$$

$$\Delta g^{ideal} = x_w \ln x_w + (1 - x_w) \ln x_s \quad (9b)$$

$$\Delta g^{excess} = x_w \ln \gamma_w + (1 - x_w) \ln \gamma_s, \quad (9c)$$

where Δg^{mix} is the normalized change in Gibbs free energy of the mixture, Δg^{ideal} is the change in ideal Gibbs free energy of the mixture (Raoult's law) and Δg^{excess} is the excess Gibbs free energy of mixing quantifying the deviation from Raoult's law. In highly non-ideal solutions liquid-liquid phase separation may occur. Two compositions x_a and x_b define the water mole fraction of the two co-existing phases. Computationally, x_a and x_b can be obtained from Δg^{mix} using the area method (Eubank et al., 1992). Briefly, the state space is evaluated by computing the following area for all possible combinations x_I and x_{II}

$$A(x_I, x_{II}) = \left|\left[\Delta g^{mix}(x_{II}) + \Delta g^{mix}(x_I)\right]\left[\frac{x_{II}-x_I}{2}\right]\right| - \left|\int_{x_I}^{x_{II}} \Delta g^{mix}(x)dx\right|. \quad (10a)$$

Phase boundaries x_a and x_b exist if condition

$$A(x_a, x_b) = \max A(x_I, x_{II}); A > 0 \quad (10b)$$

is satisfied. If multiple phases coexist in phase equilibrium, the Gibbs-Duhem relationship dictates that the chemical potential of each component is equal in all phases. Therefore the water activity inside the miscibility gap is constant and the values entering Eq. 4 are subject to the constraint

$$a_w = \begin{cases} a_w(x_a) = a_w(x_b) & \text{for } x_a \le x_w \le x_b \\ \gamma_w x_w & \text{else} \end{cases}. \quad (11)$$

We note that Eubank et al. (1992) algorithm can be extended to *n*-components. Other numerically efficient approaches to find phase equilibrium, including those of *n*-component mixtures, are available in the literature (e.g. Amundson et al., 2005, 2007, Zuend et al., 2010). Comparison for phase boundaries (x_a, x_b) calculated using standard UNIFAC parameters and the Eubank method used in this model, and standard UNIFAC parameter and the algorithm in the UHAERO model (Amundson et al., 2007) are in good agreement and summarized in the supplementary information.

Model Implementation

The model was implemented to run on a personal computer using the commercial MATLAB environment (MathWorks, Inc.). Alternatively, the code runs under the Octave environment,

which is available as free software under the GNU General Public License. Correct implementation of the UNIFAC model was confirmed by comparing results from test mixtures against output from existing implementations which is further described in the supplementary information. A compound is defined by specifying a count of subgroups comprising the molecule. Eqs. (6)-(8) are solved to find γ_w for n linearly spaced values within the domain $x_w \in [0.0001, 0.9999]$. Resulting γ_w are parsed through Eqs. (9)-(11) to find the number of stable phases and to define a_w over the entire domain. These a_w are interpolated onto a higher resolution linearly gridded domain (m points) to improve the accuracy of the computation of s_c using Eq. (4). Values for n and m are selected to balance computational speed and solution accuracy. Equations (6)-(8) have linear time complexity. Equations (9)-(11) have quadratic time complexity. Thus the two algorithms have order O(n) and O(n^2), respectively. For $n > 200$, the overall model time complexity is O(n^2). For $n >$ ~800 and $m = 10000$, the resolution is sufficiently high so that the computed s_c becomes independent of the choice of n. All computations in this work were carried out for $n = 1000$ and $m = 10000$. Total model execution times for a single compound on an Intel(R) Core(TM) i7-2600 3.4 GHz microprocessor using a single core were 39 s with MATLAB version R2013a (8.1.0.604) 64-bit and 282 s with GNU Octave version 3.8.1 configured for 64-bit.

Hygroscopicity Parameter

Equation (4) is solved to find s_c for a specified dry diameter, fixed $T = 298.15$ K and $\sigma_{s/a} = 0.072$ J m^{-2}. The result is expressed in terms of the hygroscopicity parameter κ (Petters and Kreidenweis, 2007) that is defined via

$$s_c = \left\{ \max_{D \in [D_d, \infty]} \left[\frac{D^3 - D_d^3}{D^3 - D_d^3(1-\kappa)} \exp\left(\frac{4\sigma_{s/a}}{\rho_w RTD} \right) \right] - 1 \right\} \cdot 100\% \quad (12)$$

The hygroscopicity parameter is obtained by iteratively seeking the κ value that satisfies Eq. (12) for a given D_d, s_c pair. Kappa values obtained by fitting a D_d, s_c pair to Eq. (12) with the assumed temperature and surface tension conceptually correspond to "apparent hygroscopicity at standard state" (Christensen and Petters, 2012). All values in this work are apparent κ's. For simplicity

these are denoted as κ without further qualification. Observations against which the model is evaluated are summarized in the supplementary information and will be discussed further in Section 3.

Model Assumptions and Limitations

The model approach presented here is limited to liquid organic compounds. This assumption is implied in both molar volume and UNIFAC activity coefficient calculations. Comparison with observational CCN data where the reference phase state may be crystalline should be interpreted with caution. For example, CCN experiments performed with crystalline dicarboxylic acids demonstrate that for some compounds deliquescence, i.e. a solubility-controlled phase transition, must precede droplet activation (Petters and Kreidenweis, 2008). The UNIFAC approach is unable to accurately predict the solubility of these compounds if they existed in their crystalline solid state. If, however, the compound is in metastable aqueous solution, the UNIFAC prediction is expected to be valid to within the general accuracy of the specific model implementation. Under atmospheric conditions where the organic compounds are embedded in a matrix comprising a multitude of organic compounds, liquid or amorphous solid is the prevailing stable phase (Marcolli et al., 2004). Furthermore, since metastable states with hygroscopically bound water appear to dominate in the atmosphere (Rood et al. 1989, Nguyen et al., 2014) the liquid assumption may not be a serious limitation. Nonetheless, it is unclear whether the assumption of a liquid-like reference state is a serious limitation if the organic particles are highly viscous (Vaden et al. 2011, Shiraiwa et al. 2011, Zobrist et al. 2011, Renbaum-Wolff et al. 2013).

Other limitations of the UNIFAC method are the problems of accounting for group proximity effects and the inability to distinguish between isomers. Proximity effects occur when polar groups are separated by less than three to four carbon atoms (Topping et al., 2005). Since only the number of groups of type *i* are specified, all isomers are modeled to have identical κ values. Although experiments show that the location of the functional group has a small and systematic effects on the observed κ (Suda et al., 2014), those effects are relatively small and beyond the resolution of the model presented here.

The application of Eq. (4) assumes that the surface tension is that of pure water. Many organic compounds found in ambient organic aerosol lower the surface tension at the solution/air interface (Tuckermann and Cammenga, 2004, Tuckerman, 2007). However, several studies have demonstrated via experiment and theory that surfactant partitioning between the bulk solution and the Gibbs surface phase greatly diminishes the effect one would predict by applying macroscopic surface tensions in Köhler theory (Li et al., 1998, Rood and Williams, 2001, Sorjamaa et al., 2004, Prisle et al., 2011). Neglecting to account for reduced surface tension and using water activity to estimate CCN activity results in an underestimate of κ of ~30% for the strong surfactant sodium dodecyl sulfate (Petters and Kreidenweis, 2013). We note that estimates of surface tension reduction for pure organic liquids can be obtained from critical pressure and boiling point (Sastri and Rao, 1994) and the Sprow and Prausnitz (1966) expression coupled with UNIFAC activity coefficients (Topping et al., 2005, Rafati et al., 2011). Combined with predictions of critical properties from functional group data (Joback and Reid, 1987), predicted binary surface tensions could be obtained for each compound. Including surfactant partitioning in Eq. (4) is possible using the expressions in Petters and Kreidernweis (2013) or similar approaches (Sorjamaa et al., 2004, Raatikainen and Laaksonen, 2011). Thorough validation against experimental data, including measurements of surface tension and CCN activity are needed before this approach should be adopted.

Relationship to other thermodynamic models and application to multicomponent systems

The basic model functionality described here can also be obtained by appropriately initializing other multicomponent equilibrium models (Ming and Russell, 2002, Raatikainen and Laaksonen, 2005, Topping et al., 2005, Clegg and Seinfeld, 2006, Amundson et al., 2007, Zuend et al., 2008) with a set of binary water/organic solutions, parsing the output through a phase equilibrium module (if not included in the thermodynamic model itself) and the Köhler model. The predicted CCN activity mostly depends on the underlying set of group interaction parameters. The output should match with the solution presented here if the same interaction parameter matrix is used. The main conceptual distinction between the approach proposed here and the approach employed by the more complex multicomponent models is our focus on

predictions for binary organic/water solutions and limitation of the scope to a narrow range of water activities relevant to CCN activation only. Accurate representation of hygroscopic growth at $a_w < \sim 0.99$ is not required and would be of secondary concern when tuning interaction parameters.

We envision that the proposed specialized model approach can be used to categorize individual compound into three miscibility regimes, analogous to the solubility regimes defined in Petters and Kreidenweis (2008). Regime I: the compound is CCN inactive and can be effectively modeled as $\kappa = 0$. Regime II: the compound is CCN active without any additional phase constraints. In turn κ is mostly determined by molar volume and slightly modulated by activity coefficients. Regime III: the compounds' CCN activity is limited due to miscibility constraints. In turn κ is highly sensitive to overall water content and can either have $\kappa \sim 0$ or express κ according to its molar volume. Once pure component κ's are predicted and stored in a database, the overall OA κ in mixed particles can be calculated quickly using the volume weighted mixing rule (Petters and Kreidenweis, 2007). This compound-by-compound treatment of multicomponent mixtures assumes that solute-solute interactions are negligible. Salting-in and salting-out of solution effects are not captured. Effective κ-values for compounds falling into the limited miscibility regime may be misrepresented in this treatment. Whether such effects are important will depend on the fraction of compounds in a mixture that fall into the limited miscibility regime and whether the proposed approach of intermediate complexity – modelling binary solutions coupled with a linear mixing rule – ultimately proves sufficiently accurate model the evolution of ambient OA. In the following we use experimental data to demonstrate that the outlined UNIFAC model is suitable to categorize compounds into these three regimes.

Results and Discussion

Experimental data for validation was compiled from the literature. A detailed summary of the compound names, chemical structures, physicochemical properties, CCN observations, and observed κ_{app}'s is provided in the supplementary information (Tables S3-S7). This set features compounds with mostly linear carbon backbones C_4 to C_{18} and O:C ratio between 0.1 and 1. The data are grouped into model compounds for primary organic aerosol (POA, Table S3),

functionalized hydroperoxy ethers (Table S4), hydroxynitrates (Table S5), carboxylic acids (Table S6), and carbohydrates (Table S7). Compounds included in Table S3 are long chain molecules that have hydrophobic tails (> 14 methylene groups) and a single terminal carboxyl or hydroxyl group. Representative example compounds are oleic acid or cetyl alcohol. Compounds in Table S4 are C_{14} functionalized hydroperoxy ethers that have 10-12 methylene groups, at least one hydroperoxide and ether group, and a second carbonyl, hydroperxide, or carboxyl group. Compounds in Table S5 are functionalized hydroxynitrates featuring C_{10} to C_{15} carbon backbones with 1-3 hydroxyl and 1-4 nitrate groups. Compounds in Table S6 are C_4-C_{10} carboxylic acids that have 1-2 carboxyl and up to one carbonyl group attached to the carbon backbone. Finally, compounds in Table S7 are C_4-C_{18} carbohydrates that have hydroxyl groups approximately equal to the number of carbon atoms. Data in Table S3 are taken from Raymond and Pandis (2002) and Shilling et al., (2007). Data in Tables S4 and S5 are taken from the supplement of Suda et al. (2014). Data in Tables S6 and S7 are from various sources are summarized in the supplement of Petters et al. (2009b), which was updated with new compounds from Christensen and Petters (2012), and data were re-screened for quality. The compounds were selected to provide systematic variation in the number and type of functional groups with otherwise similar structure, i.e. linear or weakly branched alkane backbone with variable carbon chain length.

To illustrate model initialization and model output two example compounds from the supplementary information C_{12} dihydroxynitrate and C_{13} trihydroxynitrate, are presented in Table 1. For some of the compounds density and solubility data are available and those data are included in the supplementary information. Table 1 shows how the molecular structure is decomposed into the subgroups understood by the UNIFAC and Girolami (1994) model framework. Detailed model output for the two example compounds is illustrated in Figure 1. The predicted mole fraction dependence of Δg_{mix} suggests that the C_{13} trihydroxynitrate is miscible with water in all proportions while the C_{12} dihydroxynitrate is not. The dashed black line connecting x_a and x_b encloses the maximum positive area with the Δg_{mix} line and defines the two-phase region. Water activity derived from Δg^{mix} is graphed in the middle panel. It shows that the miscibility gap for the C_{12} dihydroxynitrate occurs at water activity close to unity. Phase gaps at

water activity near unity may result in miscibility-controlled cloud droplet activation (Petters et al., 2006), which is analogous to solubility/deliquescence limited cloud droplet activation (Shulman et al., 1998, Hori et al., 2003, Bilde and Svenningsson et al., 2004, Kreidenweis et al., 2006, Petters and Kreidenweis, 2008). Köhler curves in the right panel demonstrate miscibility-limited activation behavior. For the C_{13} trihydroxynitrate, the Köhler curve is smooth and exhibits a single maximum corresponding to the model critical supersaturation. For the C_{12} dihydroxynitrate two maxima appear. The first maximum corresponds to the point of incipient phase separation x_a. The height of the miscibility barrier depends on the dry diameter. For large dry particles where the Kelvin term does not play a significant role, the supersaturation of point x_a is reduced and the second classical Köhler maximum will control droplet activation. Similar complex Köhler curves have been reported previously (e.g. Bilde and Svenningsson, 2004, Petters and Kreidenweis, 2008). Experiments with pure crystalline sparingly-soluble organic compounds have demonstrated convincingly that the larger maximum indeed controls cloud droplet activation for solubility-limited cases (Hori et al., 2003, Bilde and Svenningsson, 2004, Hings et al., 2008). The s_c vs. D_d relationship for phase-controlled activation does not result in κ_{app} that is independent with respect to D_d (Petters and Kreidenweis, 2008). Therefore for compounds having $\kappa < \sim 0.06$ where phase separation might play a role, the observed s_c, D_d pair is included in the data tables (Table 1, Tables S3-S7) and κ values are computed from the observation and the model (Eq. 12) at the same D_d. Note that the D_d-dependent κ only plays a role in a narrow range of miscibilities. Sufficiently soluble and truly insoluble substances are not affected. In summary, Table 1 and Figure 1 demonstrate model input, illustrate model mechanics, and identify model outputs.

How well do data-derived and model-derived κ_{app} compare? For numerical comparison both κ's are included in Tables S3-S7. A graphical illustration of these is presented in Figure 2. To improve clarity compounds with predicted and modelled $\kappa < 0.001$ are clustered in the lower left corner. Such low κ's correspond to compounds that are effectively CCN inactive. The range between $\kappa = 10^{-3}$ and 10^{-5} spans a narrow range in the s_c-D_d-κ state space that characterizes CCN activity (cf. Figure 1 Petters and Kreidenweis, 2007). Resolving these differences is not

385 particularly meaningful for the organic dominated particles that typically have D_d < 300 nm.
386 Furthermore, the κ of an internally mixed particle is approximately the weighted volume fraction
387 in the mixture. For $\kappa < 10^{-3}$ the contribution to a mixed particle's κ is insensitive to the exact
388 value. Finally, although state-of-the-science size-resolved CCN measurements can resolve
389 differences in $\kappa < 10^{-3}$, compound impurities can interfere. A 1% impurity having κ similar to
390 ammonium sulphate would contribute ~0.06 to a measured particle κ. In addition, solvent
391 residuals (Huff Hartz et al., 2006, Shilling et al. 2007, Rissman et al. 2007) and control over the
392 dry particle phase state (Raymond and Pandis, 2002, Hori et al., 2003, Broekhuizen et al., 2004,
393 Bilde and Svenningson, 2004) can disproportionally bias the characterization of low κ's.
394 Combined these points justify the definition of $\kappa < 0.001$ as effectively CCN inactive.
395 Compounds in the CCN inactive corner include all compounds from Table S3, the C_{14} and C_{15}
396 hydroxnitrate, and the C_{14}-trinitrate. These compounds all have 11 or more methylene groups and
397 O:C ratios between 0.11 and 0.65. CCN activity of these compounds is satisfactorily predicted by
398 the model.

399 Nine compounds are predicted to be CCN inactive but have measurements indicating
400 $0.001 > \kappa_{obs} > \sim 0.03$. These are graphed below the dashed line and include C_{14} di- and tetra-
401 nitrate, C_{13} hydroxynitrate, C_{14} and C_{15} dihydroxynitrate, the remaining hydroperoxide ethers
402 from Table S4, and cis-pinonic acid. The observed C_{14} di- and tetra-nitrate are barely larger than
403 the cutoff for CCN inactive. Variation of κ between the C_{14} di- and tri- and tetra-nitrate (cf.
404 Figure 2, Suda et al., 2014) implies that the trinitrate has lower κ than the di- and tetra-nitrate,
405 which suggests that some random variability in the data is superimposed on the trend. Similarly,
406 the observations show that the C_{14}- and C_{15} dihydroxynitrate are slightly more CCN active than
407 the C_{13} dihydroxynitrate. Although this is possible such behaviour is not plausible due to the
408 well-established hydrophobic nature of the added CH_x groups. One possible explanation for the
409 discrepancies is the sensitivity of observed κ's to trace contamination. Each of the compounds
410 was purified via high performance liquid chromatography (HPLC,Suda et al., 2014) but degree
411 of purification likely varied between compounds. Furthermore, experimental uncertainty for the
412 HPLC-CCN method used is slightly larger than for standard methods since it requires application

of fast flow scans. Finally, the data are from a single set of experiments. More data are needed before attributing the mismatch to either model or measurement error.

Another notable outlier is adipic acid. Here, the observed $\kappa < 0.01$ corresponds to the solubility-limited value that is referenced against its solid crystalline phase state. In contrast, the predicted value $\kappa=0.14$ is in good agreement with the molar volume prediction ($\kappa = 0.17$, cf. Figure 4, Christensen and Petters, 2012) and observed κ that adipic acid particles express when solubility limitations are removed (cf. Figure 1, Hings et al., 2008). This scenario was selected to illustrate the inability of the UNIFAC model to treat solid phases. It therefore cannot capture deliquescence and deliquescence/solubility limited activation. In atmospheric OA multiple organic compounds likely form an amorphous supercooled melt (Marcolli et al., 2004) and metastable aqueous solutions are ubiquitous (Rood et al., 1995). Thus the metastable prediction would be valid to account for adipic acid in the context of atmospheric OA.

A series of carboxylic acids and carbohydrates cluster near the 1:1 line at $\kappa > \sim 0.06$. These compounds are generally highly functionalized having at least 2 carboxyl, hydroxyl, or carbonyl group for every 4 carbon atoms. The O:C ratio always exceeds 0.5 and is close to 1 for many of the compounds. For the predictions, activity coefficients approach unity, compounds are miscible in water in all proportions, and model κ's closely track the prediction based on estimated molar volume. Overall comparison of predicted vs. observed κ is approximately within a factor of two and this range is similar to predictions that are based on actual molar volume (cf. Figure 2, Petters et al., 2009b).

The series of hydroxynitrates, dihydroxynitrates, and trihydroxynitrates for different carbon chain length also cluster near the 1:1 line. The spread is within approximately a factor of two and similar to that of the carboxylic acids and carbohydrates. These compounds span the entire range from $\kappa < 0.001$ to $\kappa \sim 0.1$ and have as few as two hydroxyl and one nitrate group per 13 carbon atoms (C_{13} dihydroxynitrate). The model appears to accurately predict the influence of the methylene and hydroxyl groups on the transition from immiscible and CCN inactive to sufficiently miscible and CCN active according to the molar volume of the compound. For the C_{11}, C_{12}, and C_{13} dihydroxynitrates the predicted miscibility-limited activation demonstrated in

Figure 1 seems to adequately explain the transition. The accurate model prediction of this sensitive transition regime is encouraging, especially since no adjustment was made to the a_{mn} group interaction parameters for [OH], [CH_x], and [H_2O] groups.

In summary, Figure 2 demonstrates four capabilities of the model. First, the model has good skill in correctly classifying effectively CCN inactive compounds ($\kappa < 0.001$). Second, the model captures the molar volume dependent activation of highly functionalized compounds (low molecular weight dicarboxylic acids and polysaccharides). Scatter between predicted and observed κ is approximately within a factor of two and considered acceptable taking into account the considerable diversity in the underlying CCN data. We note that uncertainties in molar volume estimation of $v_s \sim \pm 10\%$ stemming from the Girolami et al. (1994) method correspond to ±10% error in predicted κ for these compounds, which is significantly less than the observed scatter in the data (Petters et al., 2009). Third, the model predicts that miscibility limitations are the cause for poor CCN activity of weakly functionalized hydrocarbons, and the phase separation information can be used to quantitatively predict the transition between sufficiently miscible and effectively immiscible species. Finally, the model seems to accurately capture the main functional group dependencies observed previously (Suda et al., 2014): a strong promoting effect of hydroxyl, a weak promoting effect for hydroperoxides, a negligible or inhibiting effect of nitrate, and inhibiting effect of methylene groups on CCN activity. How, then, can one quantify the model sensitivity of κ to the addition of functional groups to otherwise similar molecules?

Simulation of homologous series can be used to derive these sensitivities. Figure 3 shows modelled κ's for a series of functionalized *n*-alkanes. The gradual decreasing trend of κ with increasing carbon number is due to the increase in molar volume. A steep decline is observed when a critical carbon number is exceeded. Beyond this point the additional methylene groups reduce the miscibility with water and render the compound effectively CCN inactive. For example, CCN activity for a C_{16} trihydroxyalkane is controlled mostly by molar volume while C_{18} trihydroxyalkane is effectively CCN inactive. The critical carbon number is C_7, C_{12}, C_{16}, C_{20}, and C_{24} for the mono-, di-, tri-, tetra-, and penta-hydroxyalkanes, respectively. Starting with an *n*-alkane, the most dramatic effect of adding functional groups is to render the molecule miscible

with water. Contrasting the critical carbon number for different homologous series is can be used as a measure of a particular groups' ability to transform the molecule such that it is sufficiently miscible in water and can express its molar volume κ. The hydroxyalkane series shows that approximately one hydroxyl group is needed to compensate for the addition of 4 methylene groups (i.e. to maintain miscibility at the composition of the critical carbon number). Expressed as a ratio, $\Delta[CH_n]/\Delta[OH] \sim -4/1$. Similar ratios for the other groups are derived from the shifts in the dihydroxyalkane series upon further functionalization: $\Delta[CH_n]/\Delta[C(=O)OH] \sim -5/2$, $\Delta[CH_n]/\Delta[CH_nC(=O)] \sim -2/3$, $\Delta[CH_n]/\Delta[HC(=O)] \sim -4/2$, $\Delta[CH_n]/\Delta[CH_n(O)] \sim -2/4$, $\Delta[CH_n]/\Delta[CH_n(OOH)] \sim -2/2$, and $\Delta[CH_n]/\Delta[CH_nC(=O)] \sim -2/3$, and $\Delta[CH_n]/\Delta[CH_nONO_2] \sim 2/3$. This leads to a sorting of relative effectiveness of the groups in promoting miscibility, hydroxyl (-4) > acid (-2.5) > aldehyde (-2) > hydroperoxide (-1) > carbonyl (-0.66) > ether (-0.5) > nitrate (0.66), where the number in parentheses corresponds to the $\Delta[CH_n]/\Delta[n]$. According to this model the addition of nitrate groups is in the same direction as methylene groups, i.e. it reduces miscibility. This finding is consistent with CCN experiments on alkenes reacted with NO_3 radicals (Suda et al., 2014, supplement), and the known low miscibility of organic nitrates in water (Boschan et al., 1995). Furthermore sorting of the different functional groups is qualitatively consistent with the sensitivity of κ to the addition of functional groups derived from CCN data (Table S5, Suda et al., 2014).

Treatment of OA evolution in the atmosphere

The computational speed of the model is relatively slow. The slow speed is due to the need to evaluate the entire range of mole fractions in order to determine the phase boundaries. Improvement in model execution speed is likely possible via algorithm optimization. Furthermore, parallel execution of the code is possible. With a regular workstation it is feasible to perform offline computation of ~10^6 κ's for a large set of compounds produced by the Generator of Explicit Chemistry and Kinetics of Organics in the Atmosphere (GECKO-A) or similar models. Once pure component κ's are predicted, the evolution of the overall OA κ in mixed particles can be calculated quickly using the linear mixing rule (Petters and Kreidenweis, 2007), subject to the limitations of this approach discussed in Section 2. One additional

limitation is the need for algorithms that automatically map the computer-generated simplified molecular-input line-entry system (SMILES) structures (e.g. Table 3, Lee-Taylor et al., 2015) to UNIFAC groups. Several of these structures are bridged and even manual mapping of those structures to UNIFAC groupings will necessitate definition of new groups with unknown volume, surface, and interaction parameters. Separate studies are needed to establish the minimal number of new groups that would be needed to obtain optimal coverage for the set of compounds of interest.

Summary and Conclusions

This paper describes how functional group contribution methods can be used to estimate the CCN activity of pure organic compounds. Group interaction parameters were taken from a mix of sources and used without further tuning. Model fidelity was evaluated against a database of published CCN data. Weakly functionalized alkanes are correctly classified as effectively CCN inactive (defined as $\kappa < 0.001$). Highly functionalized and water-soluble molecules are predicted to activate in accordance with the estimated molar volume and generally predictions agree with observations within a factor of two. Liquid-liquid phase separation is predicted to occur for compounds with few functional groups and phase separation is predicted to control κ. The model adequately reproduces the observation that hydroxyl groups strongly promote CCN activity while nitrate groups inhibit CCN activity. A few outliers in the model evaluation may be explained by the combination of CCN measurement uncertainty, compound purity, uncertainty in dry particle phase state, and insufficiently tuned group interaction parameters. However, more systematic data on weakly functionalized compounds, including repeat studies, are needed before a retuning of parameters is justified. The model makes new predictions about the relative effectiveness of the groups in promoting miscibility. Most notably, it predicts that hydroperoxides have much less of an effect than hydroxyl, which is slightly surprising since one would expect the hydrogen bonding to be similar. The model state space can serve as a rough guide to define test conditions to quantify via experiment the effectiveness of adding one or more functional groups to a carbon backbone.

Although this work is limited to a few functional groups, the presented framework is general since interaction parameters are available for a wide range of groups. For atmospheric purposes, amines, olefins, and aromatic compounds are the most relevant groups that need to be added. Few, if any systematic CCN data for these groups are available. However, the success of the current model to estimate κ without the need to tune parameters could be taken as indication that first order predictions can be obtained until such data become available.

Code Availability

Source code and example scripts demonstrating model initialization for the compounds presented in this study is available as supplementary information to this manuscript.

Acknowledgements

This work was funded by the Department of Energy, Office of Biological and Environmental Sciences under grant DE-SC0006633.

References

Amundson, N. R., Caboussat, A., He, J. W., and Seinfeld, J. H.: An optimization problem related to the modeling of atmospheric organic aerosols, C. R. Acad. Sci. Paris, Ser. I, 340, 765–768, doi:10.1016/j.crma.2005.04.018, 2005.

Amundson, N. R., Caboussat, A., He, J. W., Martynenko, A. V., Landry, C., Tong, C., and Seinfeld, J. H.: A new atmospheric aerosol phase equilibrium model (UHAERO): organic systems, Atmos. Chem. Phys., 7, 4675–4698, doi:10.5194/acp-7-4675-2007, 2007.

Andreae, M. O. and Rosenfeld, D.: Aerosol–cloud–precipitation interactions, Part 1. The nature and sources of cloud-active aerosols, Earth-Sci. Rev., 89, 13–41, doi:10.1016/j.earscirev.2008.03.001, 2008.

Barley, M. H., Topping, D. O., and McFiggans, G.: The critical assessment of liquid density estimation methods for multifunctional organic compounds, J. Phys. Chem. A, 15, 117(16). 3428–3441, doi:10.1021/jp304547r, 2013.

Bilde, M. and Svenningsson, B.: CCN activation of slightly soluble organics: The importance of small amounts of inorganic salt and particle phase, Tellus, 56B, doi:10.1111/j.1600-0889.2004.00090.x, 128–134, 2004

Bondi, A.: van der Waals Volumes and Radii, J. Phys. Chem., 68(3), 441-451, doi:10.1021/j100785a001, 1964.

Boschan, R., Merrow, R. T., and van Dolah, R. W.: The chemistry of nitrate esters. Chem. Rev., 55(3), 485-510, DOI 10.1021/cr50003a001, 1955.

Broekhuizen, K., Kumar, P. P., and Abbatt, J. P. D.: Partially soluble organics as cloud condensation nuclei: Role of trace soluble and surface active species, Geophys. Res. Lett., 31(1), L01107, doi:10.1029/2003GL018203, 2004.

Cappa, C. D., Che, D. L., Kessler, S. H., Kroll, J. H., and Wilson, K. R.: Variations in organic aerosol optical and hygroscopic properties upon heterogeneous OH oxidation, J. Geophys. Res., 116, D15204, doi:10.1029/2011JD015918, 2011.

Chang, R. Y.-W., Slowik, J. G., Shantz, N. C., Vlasenko, A., Liggio, J., Sjostedt, S. J., Leaitch, W. R., and Abbatt, J. P. D.: The hygroscopicity parameter (κ) of ambient organic aerosol at a field site subject to biogenic and anthropogenic influences: relationship to degree of aerosol oxidation, Atmos. Chem. Phys., 10, 5047-5064, doi:10.5194/acp-10-5047-2010, 2010.

Choi, M. Y. and Chan, C. K.: Continuous measurements of the water activities of aqueous droplets of water-soluble organic compounds, J. Phys. Chem. A., 106, 4566–4572, doi:10.1021/jp013875o, 2002.

Christensen, S. I. and Petters, M. D.: The role of temperature in cloud droplet activation, J. Phys. Chem. A., 116(39), 9706-9717, doi:10.1021/jp3064454, 2012.

574 Clegg, S. L., and Seinfeld, J. H.: Thermodynamic models of aqueous solutions containing
575 inorganic electrolytes and dicarboxylic acids at 298.15 K. I., The acids as non-dissociating
576 components, J. Phys. Chem. A, 110, 5692–5717, doi:10.1021/jp056150j, 2006.

577 Compernolle, S., Ceulemans, K., and Müller, J.-F.: Influence of non-ideality on condensation to
578 aerosol, Atmos. Chem. Phys., 9, 1325-1337, doi:10.5194/acp-9-1325-2009, 2009.

579 Duplissy, J., DeCarlo, P. F., Dommen, J., Alfarra, M. R., Metzger, A., Barmpadimos, I., Prevot,
580 A. S. H., Weingartner, E., Tritscher, T., Gysel, M., Aiken, A. C., Jimenez, J. L., Canagaratna, M.
581 R., Worsnop, D. R., Collins, D. R., Tomlinson, J., and Baltensperger, U.: Relating hygroscopicity
582 and composition of organic aerosol particulate matter, Atmos. Chem. Phys., 11, 1155-1165,
583 doi:10.5194/acp-11-1155-2011, 2011.

584 Eubank, P. T., Elhassan, A. E., Barrufet, M. A., and Whiting, W. B.: Area method for prediction
585 of fluid-phase equilibria, Ind. Eng. Chem. Res., 31, 942-949, doi:10.1021/ie00003a041,1992.

586 Fredenslund, A., Jones, R. L., and Prausnitz, J. M.: Group-contribution estimation of activity-
587 coefficients in nonideal liquid-mixtures, AIChE J., 21, 1086–1099, doi:10.1002/aic.690210607,
588 1975.

589 George, I. J. and Abbatt, J. P. D.: Heterogeneous oxidation of atmospheric aerosol particles by
590 gas-phase radicals, Nature Chem., 2, 713–722, doi:10.1038/nchem.806, 2010.

591 Girolami G. S.: A Simple "Back of the Envelope" Method for Estimating the Densities and
592 Molecular Volume of Liquids and Volumes, J. Chem. Education, 71(11), 962–964, doi:
593 10.1021/ed071p962, 1994.

594 Hansen, H. K., Rasmussen, P., Fredenslund, A., Schiller, M., and Gmehling, J.: Vapor-liquid
595 equilibria by UNIFAC group contribution. 5. Revision and extension, Ing. Eng. Chem. Res.,
596 30(10), 2352-2355, doi:10.1021/ie00058a017, 1991.

597 Heald, C. L., Henze, D. K., Horowitz, L. W., Feddema, J., Lamarque, J. F., Guenther, A., Hess, P.
598 G., Vitt, F., Seinfeld, J. H., Goldstein, A. H., and Fung, I.: Predicted change in global secondary
599 organic aerosol concentrations in response to future climate, emissions, and land use change, J.
600 Geophys. Res., 113, D05211, doi:10.1029/2007jd009092, 2008.

601 Hings, S. S., Wrobel, W. C., Cross, E. S., Worsnop, D. R., Davidovits, P., and Onasch, T. B.: CCN
602 activation experiments with adipic acid: effect of particle phase and adipic acid coatings on soluble
603 and insoluble particles, Atmos. Chem. Phys., 8, 3735-3748, doi:10.5194/acp-8-3735-2008, 2008.

604 Hori, M., Ohta, S., Murao, N., and Yamagata, S.: Activation capability of water soluble organic
605 substances as CCN, J. Aerosol. Sci., 34, 419–448, doi:10.1016/S0021-8502(02)00190-8, 2003.

606 Huff Hartz, K. E., Tischuk, J. E., Chan, M. N., Chan, C. K., Donahue, N. M., and Pandis, S. N.:
607 Cloud condensation nuclei activation of limited solubility organic aerosol, Atmos. Environ., 40,
608 605–617, doi:10.1016/j.atmosenv.2005.09.076, 2006.

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

620 Joback K. G. and Reid R. C.: Estimation of pure-component properties from group-contributions,
621 Chem. Eng. Commun., 57, 233–243, doi:10.1080/00986448708960487, 1987.

622 Kreidenweis, S. M., Petters, M. D., and DeMott, P. J.: Deliquescence-controlled activation of or-
623 ganic aerosols, Geophys. Res. Lett., 33, L06801, doi:10.1029/2005GL024863, 2006.

624 Kroll, J. H. and Seinfeld, J. H.: Chemistry of secondary organic aerosol: Formation and evolution
625 of low-volatility organics in the atmosphere, Atmos. Environ., 42(16), 3593-3624,
626 doi:10.1016/j.atmosenv.2008.01.003, 2008.

627 Kuwata, M., Zorn, S. R., and Martin, S. T.: Using elemental ratios to predict the density of 167
628 organic material composed of carbon, hydrogen, and oxygen, Environ. Sci. Technol., 46(2), 787-
629 794, doi:10.1021/es202525q, 2012.

630 Kuwata, M., Shao, W., Lebouteiller, R., and Martin, S. T.: Classifying organic materials by
631 oxygen-to-carbon elemental ratio to predict the activation regime of Cloud Condensation Nuclei
632 (CCN), Atmos. Chem. Phys., 13, 5309-5324, doi:10.5194/acp-13-5309-2013, 2013.

633 Lambe, A. T., Onasch, T. B., Massoli, P., Croasdale, D. R., Wright, J. P., Ahern, A. T.,
634 Williams, L. R., Worsnop, D. R., Brune, W. H., and Davidovits, P.: Laboratory studies of the
635 chemical composition and cloud condensation nuclei (CCN) activity of secondary organic
636 aerosol (SOA) and oxidized primary organic aerosol (OPOA), Atmos. Chem. Phys., 11, 8913-
637 8928, doi:10.5194/acp-11-8913-2011, 2011.

638 Lee-Taylor, J., Hodzic, A., Madronich, S., Aumont, B., Camredon, M., and Valorso, R.: Multiday
639 production of condensing organic aerosol mass in urban and forest outflow, Atmos. Chem. Phys.,
640 15, 595-615, doi:10.5194/acp-15-595-2015, 2015.

641 Li, Z., Williams, A. L., and Rood, M. J.: Influence of soluble surfactant properties on the
642 activation of aerosol particles containing inorganic solute, J. Atmos. Sci., 55, 1859–1866, 1998.

643 Marcolli, C., Luo, B. P., and Peter, T.: Mixing of the organic aerosol fractions: Liquids as the
644 thermodynamically stable phases, J. Phys. Chem. A, 108, 2216–2224, doi:10.1021/jp036080l,
645 2004.

646 Marcolli, C. and Peter, Th.: Water activity in polyol/water systems: new UNIFAC
647 parameterization, Atmos. Chem. Phys., 5, 1545–1555, doi:10.5194/acp-5-1545-2005, 2005.

648 Massoli, P., Lambe, A. T., Ahern, A. T., Williams, L. R., Ehn, M., Mikkila, J., Canagaratna, M.
649 R., Brune, W. H., Onasch, T. B., Jayne, J. T., Petaja, T., Kulmala, M., Laaksonen, A., Kolb, C.
650 E., Davidovits, P., and Worsnop, D. R.: Relationship between aerosol oxidation level and
651 hygroscopic properties of laboratory generated secondary organic aerosol (SOA) particles,
652 Geophys. Res. Lett., 37, L24801, doi:10.1029/2010GL045258, 2010.

653 Mei, F., Setyan, A., Zhang, Q., and Wang, J.: CCN activity of organic aerosols observed
654 downwind of urban emissions during CARES, Atmos. Chem. Phys., 13, 12155-12169,
655 doi:10.5194/acp-13-12155-2013, 2013.

656 Ming, Y. and Russell, L. M.: Predicted hygroscopic growth of sea salt aerosol, J. Geophys. Res.,
657 106(D22), 28259–28274, doi:10.1029/2001JD000454, 2001.

658 Ming, Y. and Russell, L. M.: Thermodynamic equilibrium of organic–electrolyte mixtures in
659 aerosol particles, AIChE J., 48, 1331–1348, doi:10.1002/aic.690480619 2002.

660 Mochida, M. and Kawamura, K.: Hygroscopic properties of levoglucosan and related organic
661 compounds characteristic to biomass burning aerosol particles, J. Geophys. Res., 109, Art. No.
662 D21202, doi:10.1029/2004JD004962, 2004.

663 Moore, R. H. and Raymond, T. M.: HTDMA analysis of multicomponent dicarboxylic acid
664 aerosols with comparison to UNIFAC and ZSR, J. Geophys. Res.-Atmos., 113, D4, Art. No.
665 D04206, doi:10.1029/2007jd008660, 2008.

666 Nguyen, T. K. V., Petters, M. D., Suda, S. R., and Carlton, A. G.: Trends in particle phase liquid
667 water during the Southern Oxidant and Aerosol Study, Atmos. Chem. Phys. Discuss., 14, 7469-
668 7516, doi:10.5194/acpd-14-7469-2014, 2014.

669 Peng, C., Chan, M. N., and Chan, C. K.: The hygroscopic properties of dicarboxylic and
670 multifunctional acids: Measurements and UNIFAC predictions, Environ. Sci. Technol., 35,
671 4495–4501, doi:10.1021/es0107531, 2001.

672 Petters, M. D., Kreidenweis, S. M., Snider, J. R., Koehler, K. A., Wang, Q. and Prenni, A. J.:
673 Cloud droplet activation of polymerized organic aerosol, Tellus B, 58, 195-205,
674 doi:10.1111/j.1600-0889.2006.00181.x, 2006.

675 Petters, M. D. and Kreidenweis, S. M.: A single parameter representation of hygroscopic growth
676 and cloud condensation nucleus activity, Atmos. Chem. Phys., 7, 1961-1971, doi:10.5194/acp-7-
677 1961-2007, 2007.

678 Petters, M. D. and Kreidenweis, S. M.: A single parameter representation of hygroscopic growth
679 and cloud condensation nucleus activity – Part 2: Including solubility, Atmos. Chem. Phys., 8,
680 6273-6279, doi:10.5194/acp-8-6273-2008, 2008.

681 Petters, M. D., Wex, H., Carrico, C. M., Hallbauer, E., Massling, A., McMeeking, G. R., Poulain,
682 L., Wu, Z., Kreidenweis, S. M., and Stratmann, F.: Towards closing the gap between hygroscopic
683 growth and activation for secondary organic aerosol – Part 2: Theoretical approaches, Atmos.
684 Chem. Phys., 9, 3999–4009, doi:10.5194/acp-9-3999-2009, 2009a.

685 Petters, M. D., Kreidenweis, S. M., Prenni, A. J., Sullivan, R. C., Carrico, C. M., Koehler, K. A.,
686 and Ziemann, P. J.: Role of molecular size in cloud droplet activation, Geophys. Res Lett., 36,
687 L22801, doi:10.1029/2009GL040131, 2009b.

688 Petters, M. D. and S. M. Kreidenweis (2013), A single parameter representation of hygroscopic
689 growth and cloud condensation nucleus activity - Part 3: Including surfactant partitioning, Atmos
690 Chem Phys, 13(2), 1081-1091, doi:10.5194/acp-13-1081-2013.

691 Poulain, L., Wu, Z., Petters, M. D., Wex, H., Hallbauer, E., Wehner, B., Massling, A.,
692 Kreidenweis, S. M., and Stratmann, F.: Towards closing the gap between hygroscopic growth
693 and CCN activation for secondary organic aerosols – Part 3: Influence of the chemical
694 composition on the hygroscopic properties and volatile fractions of aerosols, Atmos. Chem.
695 Phys., 10, 3775-3785, doi:10.5194/acp-10-3775-2010, 2010.

696 Prausnitz, J. M., Lichtenthaler, R. N., and Azevedo, E. G.: Molecular thermodynamics of fluid-
697 phase equilibria, Upper Saddle River, NJ, USA, Prentice Hall PTR, 1999.

698 Prisle, N. L., Dal Maso, M., and Kokkola, H.: A simple representation of surface active organic
699 aerosol in cloud droplet formation, Atmos. Chem. Phys., 11, 4073–4083, doi:10.5194/acp-11-
700 4073-2011, 2011.

701 Raatikainen, T. and Laaksonen, A.: Application of several activity coefficient models to water-
702 organic-electrolyte aerosols of atmospheric interest, Atmos. Chem. Phys., 5, 2475-2495,
703 doi:10.5194/acp-5-2475-2005, 2005.

704 Raatikainen, T. and Laaksonen, A.: A simplified treatment of surfactant effects on cloud drop
705 activation, Geosci. Model Dev., 4, 107–116 , doi:10.5194/gmd-4-107-2011, 2011.

706 Rafati, A. A., Bagheri, A., Khanchi, A. R., Ghasemian, E., and Najafi, M.: Application of the
707 UNIFAC model for prediction of surface tension and thickness of the surface layer in the binary
708 mixtures, J. Colloid Interface Sci., 355(1), 252–258, doi: 10.1016/j.jcis.2010.12.003, 2011.

709 Raymond, T. M. and Pandis, S. N.: Cloud activation of single-component organic aerosol
710 particles, J. Geophys. Res., 107(D24), 4787, doi:4710.1029/2002JD002159, 2002.

711 Rickards, A. M. J., Miles, R. E. H., Davies, J. F., Marshall, F. H., and Reid, J. P.: Measurements
712 of the sensitivity of aerosol hygroscopicity and the k parameter to the O/C Ratio,
713 J. Phys. Chem. A, 117(51), 14120-14131, doi:10.1021/jp407991n, 2013.

714 Rissman, T. A., Varutbangkul, V., Surratt, J. D., Topping, D. O., McFiggans, G., Flagan, R. C.,
715 and Seinfeld, J. H.: Cloud condensation nucleus (CCN) behavior of organic aerosol particles
716 generated by atomization of water and methanol solutions, Atmos. Chem. Phys., 7, 2949-2971,
717 doi:10.5194/acp-7-2949-2007, 2007.

718 Renbaum-Wolff, L., Grayson, J. W., Bateman, A. P., Kuwata, M., Sellier, M., Murray, B. J.,
719 Shilling, J. E., Martin, S. T., Bertram, A. K.: Viscosity of α-pinene secondary organic material
720 and implications for particle growth and reactivity, P. Nat. Acad. Sci., vol. 110(20), 8014–8019,
721 doi:10.1073/pnas.1219548110, 2013.

722 Rood, M. J., Shaw, M. A., Larson, T. V., and Covert, D. S.: Ubiquitous nature of ambient
723 metastable aerosol, Nature, 337, 537–539, 1989.

724 Rood, M. J. and Williams, A. L.: Reply, J. Atmos. Sci., 58, 1468– 1473, 2001.

725 Sastri, S. R. S., and Rao, K. K.: A new group contribution method for predicting viscosity of
726 organic liquids, The Chem. Engineering J., 50(1), 9–25, doi:10.1016/0300-9467(92)80002-R,
727 1992.

728 Sastri, S. R. S. and Rao, K. K.: A simple method to predict surface tension of organic liquids,
729 Chem. Engin. J. Biochem. Engin. J., 59(2), 181 – 186, doi:10.1016/0923-0467(94)02946-6,
730 1995.

731 Saxena, P. and Hildemann, L. M.: Water absorption by organics: Survey of laboratory evidence
732 and evaluation of UNIFAC for estimating water activity, Environ. Sci. Technol., 31, 3318–3324,
733 doi:10.1021/es9703638, 1997.

734 Seinfeld, J. and Pandis, S.: Atmospheric chemistry and physics: from air pollution to climate
735 change, Wiley, Inc., New Jersey, USA, p. 450, 2006.

736 Shilling, J. E., King, S. M., Mochida, M., and Martin, S. T.: Mass spectral evidence that small
737 changes 574 in composition caused by oxidative aging processes alter aerosol CCN properties, J.
738 Phys. Chem., 111, 575 3358–3368, doi:10.1021/jp068822r, 2007.

739 Shiraiwa, M., Ammann, M., Koop, T., Pöschl, U.: Gas uptake and chemical aging of semisolid
740 organic aerosol particles, Proc Nat Acad Sci, 108(27), 11003–11008,
741 doi:10.1073/pnas.1103045108, 2011.

742 Shulman, M. L., Jacobson, M. C., Charlson, R. J., Synovec, R. E., and Young, T. E.: Dissolution
743 behavior and surface tension effects of organic compounds in nucleating cloud droplets, Geophys.
744 Res. Lett., 23, 277–280, doi:10.1029/95GL03810, 1996.

745 Sorjamaa, R., Svenningsson, B., Raatikainen, T., Henning, S., Bilde, M., and Laaksonen, A.: The
746 role of surfactants in Köhler theory reconsidered, Atmos. Chem. Phys., 4, 2107–2117,
747 doi:10.5194/acp-4-2107-2004, 2004.

748 Sprow, F. B. and Prausnitz, J. M.: Surface tensions of simple liquids, T. Faraday Soc., 62 (521P),
749 1097–1104, doi:10.1039/TF9666201105, 1966.

750 Suda, S. R., Petters, M. D., Yeh, K., Strollo, C., Matsunaga, A., Faulhaber, A., Ziemann, P. J.,
751 Prenni, A. J., Carrico, C. M., Sullivan, R. C., and Kreidenweis, S. M.: Influence of functional
752 groups on organic aerosol cloud condensation nucleus activity, Environ. Sci. Technol., 48,
753 10182–10190, 2014.

754 Topping, D. O., McFiggans, G. B., and Coe, H.: A curved multi-component aerosol
755 hygroscopicity model framework: Part 2 – Including organic compounds, Atmos. Chem. Phys.,
756 5, 1223–1242, doi:10.5194/acp-5-1223-2005, 2005.

757 Topping, D., Barley, M., and McFiggans G.: Including phase separation in a unified model to
758 calculate partitioning of vapours to mixed inorganic–organic aerosol particles. Faraday Discuss.,
759 165, 273–288, doi:10.1039/C3FD00047H, 2013.

760 Tuckermann, R.: Surface tension of aqueous solutions of water- soluble organic and inorganic
761 compounds, Atmos. Environ., 41, 6265–6275, 2007.

762 Tuckermann, R. and Cammenga, H. K.: The surface tension of aqueous solutions of some
763 atmospheric water-soluble organic compounds, Atmos. Environ., 38, 6135–6138, 2004.

764 Vaden, T. D., D. Imre, J. Beránek, M. Shrivastava, and A. Zelenyuk: Evaporation kinetics and
765 phase of laboratory and ambient secondary organic aerosol, Proc Nat Acad Sci, 108(6), 2190–
766 2195, doi:10.1073/pnas.1013391108, 2011.

767 Ziemann, P. J. and Atkinson, R: Kinetics, products, and mechanisms of secondary organic
768 aerosol formation, Chem. Soc. Rev., 41(19), 6582–6605, doi:10.1039/C2CS35122F, 2012.

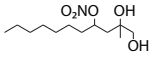
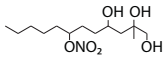
769 Zobrist, B., Soonsin, V., Luo, B. P., Krieger, U. K., Marcolli, C., Peter, T., and Koop, T.: Ultra-
770 slow water diffusion in aqueous sucrose glasses, Phys. Chem. Chem. Phys., 13, 3514–3526,
771 http://dx.doi.org/10.1039/c0cp01273ddoi:10.1039/c0cp01273d, 2011.

772 Zuend, A., Marcolli, C., Luo, B. P., and Peter, T.: A thermodynamic model of mixed organic-
773 inorganic aerosols to predict activity coefficients, Atmos. Chem. Phys., 8, 4559–4593,
774 doi:10.5194/acp-8-4559-2008, 2008.

775 Zuend, A., Marcolli, C., Peter, T., and Seinfeld, J. H.: Computation of liquid-liquid equilibria and

776 phase stabilities: implications for RH-dependent gas/particle partitioning of organic-inorganic
777 aerosols, Atmos. Chem. Phys., 10, 7795–7820, doi:10.5194/acp-10-7795-2010, 2010.
778 Zuend, A. and Seinfeld, J. H.: Modeling the gas-particle partitioning of secondary organic aerosol:
779 the importance of liquid-liquid phase separation, Atmos. Chem. Phys., 12, 3857–3882,
780 doi:10.5194/acp-12-3857-2012, 2012.

Table 1. Properties for two example chemical compounds. UNIFAC representation indicated the number and type of subgroups to represent the chemical structure MW denotes molecular weight (g mol^{-1}) and v_s denotes the model predicted molar volume (cm^3 mol^{-1}). CCN reflects the observed supersaturation and dry diameter data pair obtained from the source (Suda et al., 2014) from which observed $\boldsymbol{\kappa}$ was determined.

Name	Formula	Structure	UNIFAC representation		MW	v_s	observed CCN	Apparent κ	
			#	Subgroup			s_c (%) / D_d (nm)	observed	model
C_{12} dihydroxy nitrate	$C_{12}H_{25}O_5N$		2	CH_3	263.3	263.3	0.3 / 222	0.018	0.008
			8	CH_2					
			1	C					
			1	$CH(ONO_2)$					
			2	OH					
C_{13} trihydroxy nitrate	$C_{13}H_{27}O_6N$		2	CH_3	293.4	257.7	0.3 / 111	0.1	0.07
			8	CH_2					
			1	CH					
			1	C					
			1	$CH(ONO_2)$					
			3	OH					

Figures

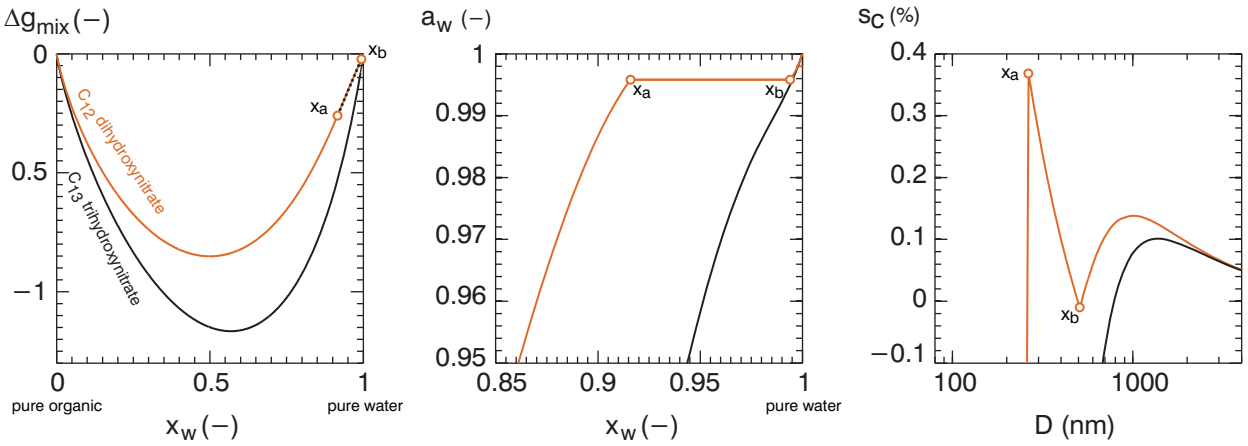


Figure 1. Modeled Δg_{mix} (left), water activity (middle), and Köhler curves (right) for C_{12} dihydroxynitrate and C_{13} trihydroxynitrate (see Table 1). Open circles denote the mole fractions x_a and x_b that correspond to the envelope of compositions where liquid-liquid phase separation is predicted for the C_{12} dihydroxynitrate.

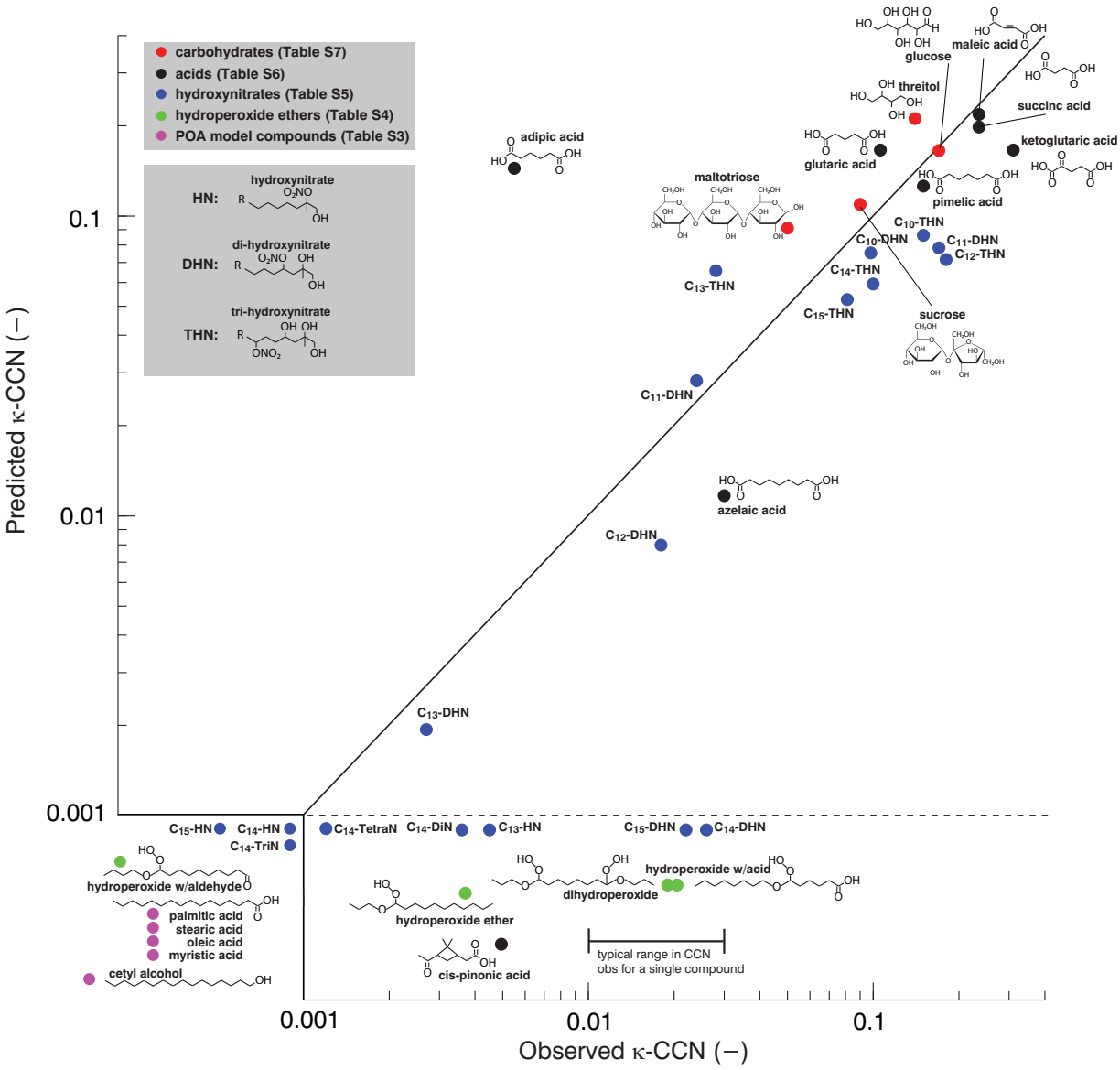
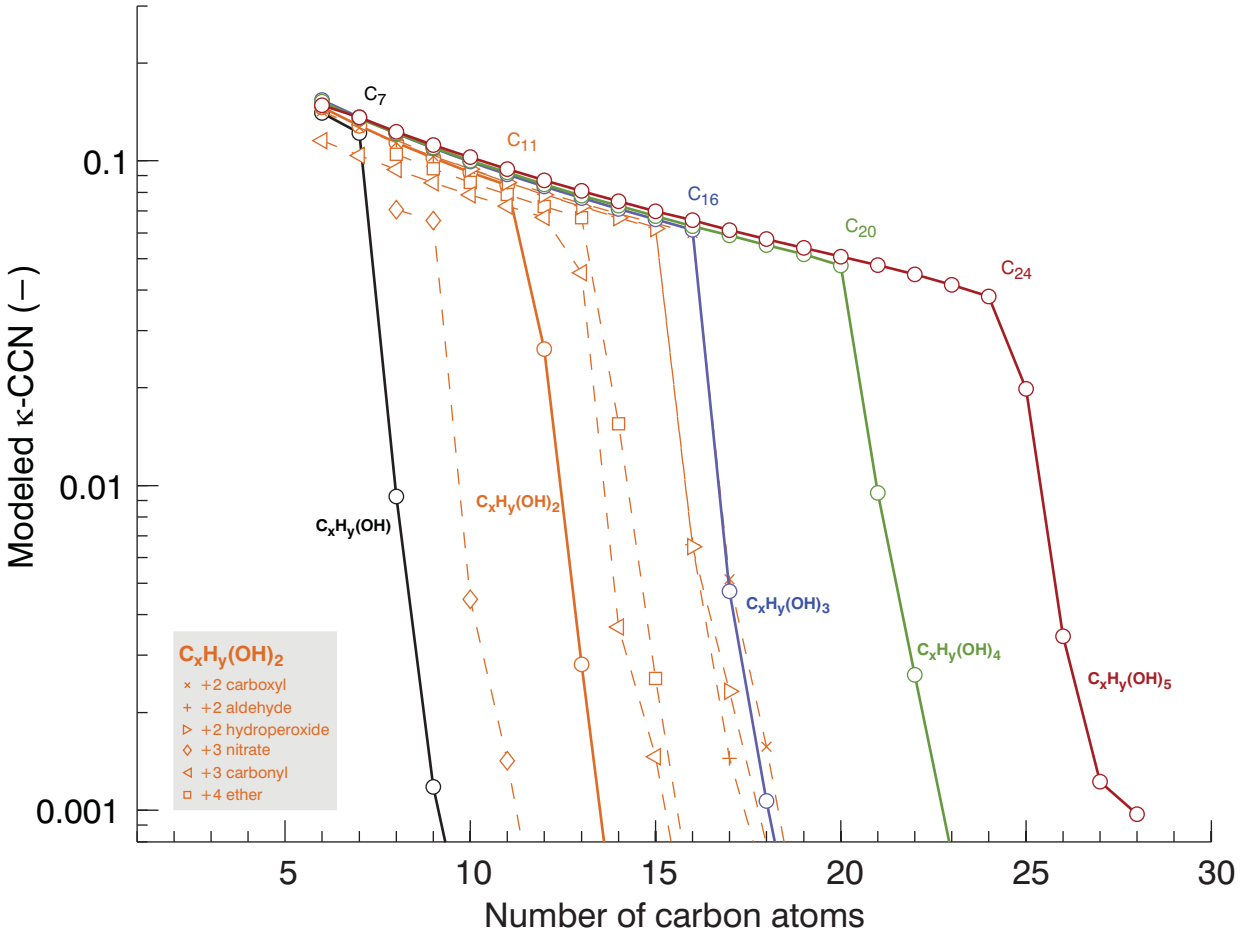


Figure 2. Model predicted vs. experimentally determined κ-values. Values $\kappa < 0.001$ are classified as CCN inactive and are clustered in the lower left corner of the graph. Colors are used to delineate the grouped source data in the supplementary information (SI). Selected structures from the SI are included in the graph. C_x-HN, C_x-DHN, and C_x-THN denote hydroxynitrate, dihydroxynitrate, and trihydroxynitrate and x denotes the total number of carbon atoms. C_{14}-DiN, C_{14}-TriN, C_{14}-TetraN denote the C_{14} dinitrate, trintrate, and tetranitrate, respectively. Points below the dashed line corresponds to compounds with predicted $\kappa < 0.001$ and observed $\kappa >$ 0.001. Typical range of observed κ CCN for peroxides is indicated by the horizontal bar.



801

802 **Figure 3.** Modeled κ values for homologous series of functionalized *n*-alkanes. Solid lines
803 correspond to alkanes with 1-5 non-terminal hydroxyl groups. Orange dashed lines correspond to
804 further functionalized dihydroxyalkanes as described in the legend. Colored carbon numbers (C_7,
805 C_{12}, C_{16}, C_{20}, and C_{24}) correspond to the largest carbon number without miscibility limited
806 activation for the respective hydroxyalkanes series.