On the implications of aerosol liquid water and phase separation for organic aerosol mass

Weiwei Hu  
*University of Colorado Boulder*

Jose L. Jimenez  
*University of Colorado Boulder*

Please see bottom of page for a full list of authors.

Follow this and additional works at: [https://scholar.colorado.edu/chem_facpapers](https://scholar.colorado.edu/chem_facpapers)
On the implications of aerosol liquid water and phase separation for organic aerosol mass

Havala O. T. Pye\textsuperscript{1}, Benjamin N. Murphy\textsuperscript{1}, Lu Xu\textsuperscript{2}, Nga L. Ng\textsuperscript{2,3}, Annmarie G. Carlton\textsuperscript{4,a}, Hongyu Guo\textsuperscript{3}, Rodney Weber\textsuperscript{3}, Petros Vasilakos\textsuperscript{2}, K. Wyat Appel\textsuperscript{1}, Sri Hapsari Budisulistiorini\textsuperscript{5}, Jason D. Surratt\textsuperscript{5}, Athanasios Nenes\textsuperscript{2,3,6,7}, Weiwei Hu\textsuperscript{8,9}, Jose L. Jimenez\textsuperscript{8,9}, Gabriel Isaacman-VanWertz\textsuperscript{10}, Pawel K. Misztal\textsuperscript{10}, and Allen H. Goldstein\textsuperscript{10,11}

\textsuperscript{1}National Exposure Research Laboratory, US Environmental Protection Agency, Research Triangle Park, NC, USA
\textsuperscript{2}School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA, USA
\textsuperscript{3}School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA
\textsuperscript{4}Department of Environmental Sciences, Rutgers University, New Brunswick, NJ, USA
\textsuperscript{5}Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
\textsuperscript{6}Institute of Environmental Research and Sustainable Development, National Observatory of Athens, Palea Penteli, 15236, Greece
\textsuperscript{7}Institute for Chemical Engineering Sciences, Foundation for Research and Technology Hellas, Patras, Greece
\textsuperscript{8}Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA
\textsuperscript{9}Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO, USA
\textsuperscript{10}Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA USA
\textsuperscript{11}Department of Civil and Environmental Engineering, University of California, Berkeley, CA USA
\textsuperscript{a}now at: Department of Chemistry, University of California, Irvine, CA, USA

Correspondence to: Havala O. T. Pye (pye.havala@epa.gov)

Received: 10 August 2016 – Published in Atmos. Chem. Phys. Discuss.: 16 August 2016
Revised: 23 November 2016 – Accepted: 11 December 2016 – Published: 6 January 2017

Abstract. Organic compounds and liquid water are major aerosol constituents in the southeast United States (SE US). Water associated with inorganic constituents (inorganic water) can contribute to the partitioning medium for organic aerosol when relative humidities or organic matter to organic carbon (OM / OC) ratios are high such that separation relative humidities (SRH) are below the ambient relative humidity (RH). As OM / OC ratios in the SE US are often between 1.8 and 2.2, organic aerosol experiences both mixing with inorganic water and separation from it. Regional chemical transport model simulations including inorganic water (but excluding water uptake by organic compounds) in the partitioning medium for secondary organic aerosol (SOA) when RH > SRH led to increased SOA concentrations, particularly at night. Water uptake to the organic phase resulted in even greater SOA concentrations as a result of a positive feedback in which water uptake increased SOA, which further increased aerosol water and organic aerosol. Aerosol properties, such as the OM / OC and hygroscopicity parameter ($\kappa_{\text{org}}$), were captured well by the model compared with measurements during the Southern Oxidant and Aerosol Study (SOAS) 2013. Organic nitrates from monoterpene oxidation were predicted to be the least water-soluble semivolatile species in the model, but most biogenically derived semivolatile species in the Community Multiscale Air Quality (CMAQ) model were highly water soluble and expected to contribute to water-soluble organic carbon (WSOC). Organic aerosol and SOA precursors were abundant at night, but additional improvements in daytime organic aerosol are needed to close the model–measurement gap. When taking into account deviations from ideality, including both inorganic (when RH > SRH) and organic water in the organic partitioning medium reduced the mean bias in SOA for routine monitoring networks and improved model performance compared to observations from SOAS. Property updates from this work will be released in CMAQ v5.2.

Published by Copernicus Publications on behalf of the European Geosciences Union.
1 Introduction

Water is a ubiquitous component of atmospheric aerosol (Nguyen et al., 2016), which can interact with organic compounds in a number of ways to influence particulate matter (PM) mass and size, human health, and Earth’s radiative balance. While constituents such as sulfate and nitrate often drive aerosol water concentrations, inorganic and organic compounds are internally mixed under humid conditions (You et al., 2013), and hydrophilic organic compounds promote the uptake of water (Saxena et al., 1995). Uptake of water into the organic phase increases particle size, making particles more effective at interacting with radiation (Chung and Seinfeld, 2002), obscuring visibility (Lowenthal and Kumar, 2016), and forming clouds (Novakov and Penner, 1993). Water can serve as a medium for partitioning of soluble (Carlton and Turpin, 2013; Pun et al., 2002) and semivolatile (Chang and Pankow, 2010; Pankow and Chang, 2008; Seinfeld et al., 2001) gases, thus contributing to particulate matter concentrations. Once in the particle phase, organic compounds can participate in water-mediated reactions such as hydrolysis, driving isoprene epoxydiol uptake to the particle (Surratt et al., 2010; Pye et al., 2013) and loss of gas-phase organic nitrates (Fisher et al., 2016).

Organic–aerosol-water interactions have been examined in a number of laboratory and field studies, and results are mixed. The lack of a consistent relationship in laboratory work may be partially due to experimental conditions such as high mass loadings that minimize the effect of water for semivolatile systems (Pankow and Chang, 2008). Laboratory studies have observed no significant change in yield with increasing relative humidity (RH) (Edney et al., 2000; Boyd et al., 2015), enhanced yields under dry conditions (Zhang et al., 2011), and higher yields with increasing aerosol water (Wong et al., 2015) depending on the precursor, oxidant, and seed. Trends in ambient aerosol organic carbon are consistent with the trend in decreasing aerosol water in the southeast US (Nguyen et al., 2015b), and Hennigan et al. (2008) observed episodic correlations of water-soluble organic carbon and water vapor. However, Hennigan et al. (2008) found no well-defined relationship over the entire summer in Atlanta, GA, and organic aerosol was not correlated with liquid water content in Pittsburgh, PA (Griffin et al., 2003). Saxena et al. (1995) found that the presence of organic compounds suppressed aerosol water in urban locations. In the atmosphere, the relative roles of different secondary organic aerosol (SOA) species change as a function of time and space, and each species may have a different sensitivity to aerosol water.

The interaction of primary organic aerosol (POA), SOA from low-volatility and semivolatile ($C_p < 3000 \mu g m^{-3}$) compounds, SOA from aqueous pathways, and the inorganic/water-rich phase is important for the concentration of organic aerosol (OA) as partitioning is a function of the availability of an absorptive medium. Current chemical transport models, including the Community Multiscale Air Quality (CMAQ) model (Carlton et al., 2010), consider SOA to be exclusively or dominantly formed via condensation of organic compounds in the absence of water. Individual model studies have examined hydrophobic and hydrophilic SOA through semi-mechanistic algorithms and surrogate structure information. Pun et al. (2002) used a decoupled approach in which organic species partitioned only to their dominant phase (aqueous vs. organic). Griffin et al. (2003) allowed compounds to partition to both phases in varying amounts based on their properties. Jathar et al. (2016) examined the implications of water uptake to the organic phase and the effects on OA concentrations. Pun (2008) allowed organic compounds to interact with water and separate into two phases if thermodynamically favorable. None of these approaches considered mixing of the inorganic and organic phases and often required computationally intensive calculations of activity coefficients. These models accounting for aerosol-water–organic interactions are not in widespread use and have not been evaluated with recently available observations of aerosol water.

Figure 1 shows the contribution of POA and water-soluble OA (determined from water-soluble organic carbon, WSOC; Sullivan et al., 2004) to total OA as observed during the Southern Oxidant and Aerosol Study (SOAS) for June 2013 in Centreville, AL. Ambient measurements of WSOC are highly correlated with oxygenated organic aerosol (OOA) (Kondo et al., 2007), and water-soluble OA accounted for 90% of total OA on average in the southeast US during summer 2013 (Washenfelder et al., 2015). WSOC has also been proposed to contain SOA from aqueous pathways with evidence for reversible (El-Sayed et al., 2015) and irreversible (El-Sayed et al., 2016) formation. CMAQ tends to overpredict the concentration of POA by almost a factor of 2 during SOAS (Pye et al., 2015). CMAQ predicts a relatively minor role for aqueous OA with the dominant source of OA in CMAQ being dry processes (other SOA in Fig. 1).

Through a series of sensitivity simulations (outlined in Sect. 2.1), this work aims to understand if interactions of aerosol water with semivolatile compounds can resolve model–measurement discrepancies and to what degree OA predicted by models should be classified as water soluble. Semiempirical SOA in the CMAQ model (Sect. 2.2) was connected to a consistent set of properties useful for predicting atmospherically relevant behavior such as interaction with aerosol water. In cases where a specific molecular species was not already used as a surrogate, aerosol properties were linked to volatility and parent hydrocarbon (Sect. 2.3). These quantities allowed molecular weights, organic matter to organic carbon (OM/OC) ratios, Henry’s law coefficients, deposition properties, hygroscopicity ($\kappa$), phase separation (Sect. 2.4), water uptake (Sect. 2.5), and deviations from ideality (Sect. 2.7) to be predicted semiempirically and influence partitioning (Sect. 2.6). In addition, the fraction WSOC was estimated for model species (Sect. 2.8), and primary vs.
secondary organic aerosol was estimated for monitoring networks (Sect. 2.9). The property updates will be available in CMAQ v5.2, and their effects on model predictions are illustrated in Sect. 3.1. The implications of the updates for OA and particle-phase liquid water content (LWC) are examined in the context of routine monitoring networks and Southern Oxidant and Aerosol Study (SOAS) observations (Sect. 3.2).

A baseline simulation including surrogate property updates detailed in Sect. 2.3 (molecular weight, Henry's law coefficients, etc.) and three sensitivity simulations examining the implications of aerosol liquid water for SOA were conducted (Fig. 2). In the baseline simulation, POA and traditional SOA were designated hydrophobic and did not interact with aerosol water or SOA produced through aqueous pathways following common chemical transport model assumptions. Two sensitivity simulations examined the implications of aerosol water on semivolatile partitioning via increases in the partitioning medium assuming ideal mixing. In one simulation (ideal $W_i$), POA, traditional SOA, aqueous SOA, and water associated with inorganic constituents were assumed to form one ideal phase when RH was above the separation relative humidity (SRH) and to undergo liquid–liquid phase separation into organic-rich (POA and traditional SOA) and water-rich (aqueous SOA and inorganic constituents) ideal phases otherwise. When one phase was predicted to be present (SRH < RH), interactions of semivolatile organic compounds and inorganic water were assumed to be ideal. When phase separation occurred, semivolatile organic compounds did not interact with water. In the second simulation, uptake of water to the organic phase ($W_o > 0$) was predicted based on its OM/OC and $\kappa$-Köhler theory (Peters and Kreidenweis, 2007) (Sect. 2.5). Thus, ideal $W_i$ and $W_o > 0$ simulations are meant to capture the effects of inorganic and organic water under ideal conditions separately. The impacts of phase separation and water uptake to organic species along with deviations from ideality determined via an activity coefficient ($\gamma$) were simulated together in the third sensitivity simulation ($\gamma \neq 1$).

### 2.2 CMAQ organic aerosol

CMAQ v5.1 contains several types of SOA with different sensitivities to aerosol water: traditional semivolatile SOA from Odum two-product representations, nonvolatile SOA produced by volatile organic compound (VOC) reaction, heterogeneously produced SOA parameterized by an uptake coefficient, semivolatile organic nitrate SOA and its hydrolysis product, and other contributions from cloud processing and accretion/oligomerization reactions (Fig. 3, Table 1). The traditional SOA systems in CMAQ include SOA from isoprene, monoterpenes, sesquiterpenes, benzene, toluene, xylene, alkanes, and polycyclic aromatic hydrocarbons (PAHs) (Carlton et al., 2010; Pye and Poulmiot, 2012). The semivolatile SOA from these precursors is allowed to oligomerize to a nonvolatile form on a 24 h timescale (Carlton et al., 2010). Currently, low-NOx oxidation of aromatics leads to nonvolatile SOA in the traditional systems. Glyoxal (GLY), methylglyoxal, and epoxides undergo heterogeneous uptake to form SOA (Pye et al., 2013, 2015). Glyoxal SOA forms using a fixed uptake coefficient of 0.0029 (Liggio et al., 2005). Following the approach of Marais et al. (2016), methylglyoxal’s uptake coefficient was scaled to the

**Figure 1.** Contribution of POA (observed biomass burning OA, BBOA; Xu et al., 2015a), SOA, water-soluble OA (estimated as $2.1 \times$ WSOC from the Particle-into-Liquid Sampler (PiLS); Sullivan et al., 2004), and aqueous (aq.) SOA (model only) to total OA during June 2013 observed at CTR during SOAS and modeled by standard CMAQ. Insoluble OA is the difference between measured total OA and water-soluble OA. Modeled “other SOA” is formed via partitioning to a dry organic phase.
glyoxal uptake coefficient by the relative Henry’s law coefficient (resulting in an uptake coefficient of $2.6 \times 10^{-4}$) in this work. Isoprene epoxydiol (IEPOX) SOA is parameterized with an uptake coefficient calculated as a function of aerosol phase constituents, including sulfate and water assuming an acid-catalyzed mechanism (Pye et al., 2013). In this work, the IEPOX organosulfate formation rate constant was updated to $8.83 \times 10^{-3}$ M$^{-2}$ s$^{-1}$ using the ratio of 2-methyltetrol to organosulfate formation rate constants from Piletic et al. (2013) and a 2-methyltetrol rate constant of $9 \times 10^{-4}$ M$^{-2}$ s$^{-1}$ (Eddingsaas et al., 2010). This organosulfate rate constant is more aggressive (overall and relative) than predicted by Riedel et al. (2016). Overestimates of the organosulfate in the model may compensate for missing IEPOX-derived SOA species such as C$_5$-alkene triols (Surratt et al., 2010) or additional oligomers (Lopez-Hilfiker et al., 2016). In addition, the Henry’s law coefficient was updated to $3.0 \times 10^7$ M atm$^{-1}$ (Nguyen et al., 2014a), which improved model predictions of 2-methyltetrols (Supplement) and total isoprene SOA. The diffusivity of IEPOX in the particle ($D_a$, cm$^2$ s$^{-1}$) was predicted by fitting a line through the data in the work of Song et al. (2015) resulting in

$$D_a = 10^{(7.18RH-12.7)}$$

(1)

for $0 \leq RH \leq 1$. Semisolid organic aerosol ($D_a < 10^{-12}$ cm$^2$ s$^{-1}$) is not expected in the humid eastern US during summer (Pajunoja et al., 2016). SOA from long-generation high-NO$_2$/high-NO$_2$ SOA species (methacrylic acid epoxide and hydroxymethyl-methyl-$\alpha$-lactone) is relatively minor (Pye et al., 2013; Marais et al., 2016), consistent with observations from SOAS ground sites (Budisulistiorini et al., 2015). All SOA produced through heterogeneous uptake is assumed to be nonvolatile in CMAQ v5.1. SOA from isoprene and monoterpane organic nitrates is semivolatile, but the nitrate groups hydrolyze in the particle to produce SOA, which is assumed to be nonvolatile, and nitric acid (Pye et al., 2015). SOA from cloud processing is predicted to result in less than 3% of total organic aerosol in CMAQ. POA and volatility-based SOA is treated as hydrophobic by default, while aqueous and in-cloud SOA is assumed to be hydrophilic and resides in a water-rich phase (CMAQv5.1 aerosol assumptions; Table 1).

### 2.3 Updating properties of semivolatilites

#### 2.3.1 Molecular properties

For SOA systems, the molecular weight and OM/OC by mass must be specified for mass-to-molecule number unit conversions within CMAQ and to allow for post-processing of organic carbon (OC) from total SOA for comparison to observations. The number of carbons per molecule ($n_C$) is also specified for the traditional semivolatile systems to allow for oligomerization to conserve carbon (Carlton et al., 2010). Historically, in CMAQ model formulations (v5.1 and prior), the $n_C$, saturation concentration ($C_i^*$), and OM/OC were set independently with the OM/OC obtained from chamber experiments and $n_C$ set to that of the parent hydrocarbon. The molecular weight was calculated to be consistent with the number of carbons and OM/OC. The OM/OC values were not a function of volatility or peroxy radical (RO$_2$) fate. Thus, all SOA species from a given parent hydrocarbon were assumed to have the same properties (OM/OC, molecular weight, number of carbons) regardless of their volatility. When viewed in the O:C vs. $C_i^*$ space (Baker et al., 2015), this leads to some apparent contradictions such as sesquiterpene SOA being more functionalized and having a longer carbon backbone at a given vapor pressure than isoprene or monoterpane SOA. This inconsistency is also seen in the molecular weight vs. $C_i^*$ space (Fig. 4). Most SOA
constituents are expected to show that molecular weight is correlated with vapor pressure ($C_i^*$) with more functionalized species having a shallower slope than less functionalized species (Shiraiwa et al., 2014). Systems examined by Shiraiwa et al. (2014) were found to reside between a line characteristic of O : C = 0 (alkane, $C_nH_{2n+2}$) and O : C = 1 (sugar, $C_nO_nH_{2n−2}$). Sesquiterpene SOA in CMAQ v5.1 resides outside the molecular corridor bounds that correspond to O : C = 0 (OM/OC=1.17) and O : C = 1 (OM/OC = 2.3 to 2.5). The CMAQv5.1 Odum two-product isoprene SOA components imply an O : C > 1 (which is possible, but not observed by Shiraiwa et al., 2014, and infrequent in the work of Chen et al., 2015).

Structure–activity relationships or group contribution methods can be used to relate vapor pressure, molecular weight, and OM/OC (or molar O : C). Donahue et al. (2011) developed a relationship between the saturation concentration of a pure species ($C_i^*$, number of carbons per molecule, and number of oxygens per molecule ($n_O$) ignoring sulfate and nitrate for use with the 2-D volatility basis set (VBS):

$$\log_{10} C_{0,i}^* = 0.475(25 - n_C) - 2.3n_O + 0.6n_C n_O / (n_C + n_O).$$

Built into this relationship are assumptions about the functionality of semivolatile organic compounds (specifically equal alcohols and ketones with acid terminal groups), the volatility of a 25 carbon alkane ($C_i^* = 1 \mu g m^{-3}$), and how a given functional group affects volatility (from the SIMPOL model; Pankow and Asher, 2008). Note that considerable variability in atmospheric aging exists in terms of the addition of functional groups as indicated on van Krevelen diagrams (Chen et al., 2015). The number of oxygen is related to the molar O : C by

$$n_O = n_C(O : C).$$

O : C can be related to the mass-based OM/OC (Simon and Bhave, 2012):

$$O : C = \frac{12}{15} \left( \frac{OM}{OC} \right) - 14 \frac{15}{15},$$

which assumes only H, O, and C atoms and produces results consistent with aerosol mass spectrometry (AMS)-determined relationships between O : C and OM/OC (Cana-
Table 1. SOA and semivolatile organic compound (SVOC) species in CMAQ v5.1-aero6i (Carlton et al., 2010; Pye and Pouliot, 2012; Pye et al., 2013, 2015). CMAQ model species names are generally preceded by the letter A to indicate aerosol. Semivolatile surrogates have a corresponding gas-phase species whose name is preceded by the letters SV.

<table>
<thead>
<tr>
<th>Species</th>
<th>Species or production pathway description</th>
<th>Partitioning medium in CMAQ v5.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALK1</td>
<td>alkane + OH SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>ALK2</td>
<td>alkane + OH SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>BNZ1</td>
<td>benzene + OH high-NO$_x$ SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>BNZ2</td>
<td>benzene + OH high-NO$_x$ SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>BNZ3</td>
<td>benzene + OH low-NO$_x$ SOA</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>DIM</td>
<td>IEPOX-derived dimers</td>
<td>Aqueous aerosol</td>
</tr>
<tr>
<td>GLY</td>
<td>glyoxal + methylglyoxal SOA</td>
<td>Aqueous aerosol</td>
</tr>
<tr>
<td>IEOS</td>
<td>IEPOX-derived organosulfate</td>
<td>Aqueous aerosol</td>
</tr>
<tr>
<td>IETET</td>
<td>2-methyltetrols</td>
<td>Aqueous aerosol</td>
</tr>
<tr>
<td>IMGA</td>
<td>2-methylglyceric acid</td>
<td>Aqueous aerosol</td>
</tr>
<tr>
<td>IMOS</td>
<td>MPAN-derived organosulfate</td>
<td>Aqueous aerosol</td>
</tr>
<tr>
<td>ISO1</td>
<td>isoprene + OH SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>ISO2</td>
<td>isoprene + OH SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>ISO3</td>
<td>acid-catalyzed isoprene SOA*</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>ISOPNN</td>
<td>isoprene dinitrate</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>MTHYD</td>
<td>organic nitrate hydrolysis product</td>
<td>Aqueous aerosol (from dry organic aerosol parent)</td>
</tr>
<tr>
<td>MTNO3</td>
<td>monoterpene nitrte</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>OLGA</td>
<td>oligomers from anthropogenic SOA/SVOCs</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>OLGB</td>
<td>oligomers from biogenic SOA/SVOCs</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>ORGC</td>
<td>glyoxal+methylglyoxal SOA</td>
<td>Cloud droplets</td>
</tr>
<tr>
<td>PAH1</td>
<td>naphthalene + OH high-NO$_x$ SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>PAH2</td>
<td>naphthalene + OH high-NO$_x$ SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>PAH3</td>
<td>naphthalene + OH low-NO$_x$ SOA</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>SQT</td>
<td>sesquiterpene + OH, O$_3$, NO$_3$, O$_3$P SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>TOL1</td>
<td>toluene + OH high-NO$_x$ SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>TOL2</td>
<td>toluene + OH high-NO$_x$ SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>TOL3</td>
<td>toluene + OH low-NO$_x$ SOA</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>TRP1</td>
<td>monoterpene + OH, O$_3$, O$_3$P SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>TRP2</td>
<td>monoterpene + OH, O$_3$, O$_3$P SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>XYL1</td>
<td>xylene + OH high-NO$_x$ SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>XYL2</td>
<td>xylene + OH high-NO$_x$ SOA/SVOC</td>
<td>Dry organic aerosol</td>
</tr>
<tr>
<td>XYL3</td>
<td>xylene + OH low-NO$_x$ SOA</td>
<td>Dry organic aerosol</td>
</tr>
</tbody>
</table>

* AIS03 contains the sum of 2-methyltetrols and IEPOX-derived organosulfates in CMAQv5.1-aero6. It is not used in aero6 as those species are represented individually. Prior to v5.1, AIS03 was determined as an enhancement over AIS01 + AIS02 based on [H$^+$] (Carlton et al., 2010).

Garatna et al., 2015). OM / OC was the focus of this work instead of O : C since OM / OC values are directly used to post-process model output for comparison to observation network measurements of OC. In addition, OM / OC ratios are a useful quantity in reconstructing the total mass of PM and could be available routinely from the Interagency Monitoring of Protected Visual Environments (IMPROVE) network in the future using Fourier transform infrared spectroscopy (FTIR) analysis (Ruthenburg et al., 2014). The molecular weight ($\tilde{M}_i$) follows as

$$\tilde{M}_i = 12n_C \left( \frac{OM}{OC} \right).$$

Equations (2) to (5) provide four equations for six unknowns: $n_O$, $n_C$, OM / OC, $C_i^*$, and $\tilde{M}_i$, $C_i^*$ was obtained from the Odum two-product fits (Odum et al., 1996) derived from laboratory data (Carlton et al., 2010; Pye and Pouliot, 2012) and $n_C$ was set to that of the parent hydrocarbon. The OM / OC and molecular weight were then calculated. $n_O$ and O : C were not needed for CMAQ (but could be easily obtained). Pankow et al. (2015) undertook a similar exercise in which they developed surrogates for each of the CMAQ v5.0 SOA species using SIMPOL and plausible structures. Their information was used when available, and Eqs. (2) to (5) were employed otherwise. For the systems on which Pankow et al. (2015) provide information, the results based on Eqs. (2) to (5) are very similar. For SOA from the explicit later-generation precursors (such as IEPOX, isoprene dinitrates, and monoterpene nitrates), the molecular properties were already tied to a specific surrogate identity. The CMAQ SOA species representing actual compounds were not updated.
2.3.2 Deposition properties

The deposition-related properties of gases, such as their solubility, diffusivity, and reactivity, are related to molecular structure and composition. CMAQ uses a resistance in series method for dry deposition (Pleim and Ran, 2011). CMAQ v4.7 through v5.1 use adipic acid (Henry’s law coefficient, $H = 2 \times 10^8$ $\text{M} \text{ atm}^{-1}$) as a wet deposition surrogate for gas-phase semivolatile organic compounds (SVOCs). Default dry deposition of SVOCs is based on acetic acid ($H = 4.1 \times 10^5 \exp(63000(298 - T))/(298T) \text{ M} \text{ atm}^{-1}$; gas-phase diffusivity ($D_g$) = 0.0944 $\text{cm}^2 \text{s}^{-1}$; dry cuticular resistance = 1200 $\text{s m}^{-1}$; LeBas molar volume = 63 cm$^3$ mol$^{-1}$).

Hodzic et al. (2014) used the Generator of Explicit Chemistry and Kinetics of Organics in the Atmosphere (GECKO) to predict products from various SOA systems commonly represented in models. Henry’s law coefficients were then estimated based on the GROup contribution Method for Henry’s law Estimate (GROMHE) (Raventos-Duran et al., 2010). GROMHE was found to reproduce Henry’s Law coefficients for organic–water systems with a mean absolute error of about 0.3 log units compared to 0.5 for HenryWin and 0.4 for SPARCv4.2 (SPARC Performs Automated Reasoning in Chemistry; Raventos-Duran et al., 2010). For SOA systems, a strong relationship was observed between saturation concentrations and Henry’s law coefficients, with chemically aged species being less volatile, more functionalized, and more soluble than their parent hydrocarbon. Although the relationship between $H$ and $C_i^*$ was relatively robust, variability in $H$ spanned many orders of magnitude for a given $C_i^*$ bin without considering how inorganic species may modify the Henry’s law coefficient. The relationships derived by Hodzic et al. (2014) were used to predict the Henry’s law coefficients as a function of $C_i^*$ for each SVOC surrogate in equilibrium with the particle in the model. An enthalpy of solvation of 50 kJ mol$^{-1}$ was also adopted to adjust the Henry’s law coefficients for temperature. Note that although the approach used by Hodzic et al. (2014) is also a group contribution method, it potentially represents the functional groups present in CMAQ SOA species with different groups than would be assumed by Eqs. (2)–(5).

Additional properties needed for deposition include the gas-phase diffusion coefficient, which was calculated as a function of molecular weight via $D_{g,i} = 1.9(M_i)^{-2/3} \text{ cm}^2 \text{s}^{-1}$ (Schnoor, 1996), and the LeBas molar volume ($V_{\text{LeBas}}$), calculated assuming ring-opened products (Mackay et al., 1993):

$$V_{\text{LeBas}} = 14.8n_C + 7.4n_O + 3.7n_H \text{ cm}^3 \text{ mol}^{-1},$$

where the number of hydrogens, $n_H$, is calculated from the molecular weight assuming only carbon, oxygen, and hydrogen. Modifications were also made to the deposition parameters affecting H$_2$O$_2$, IEPOX, and organic nitrates to produce results consistent with Nguyen et al. (2015a) (parameters available in the Supplement).

2.4 Predicting phase separation

The solubility of an organic compound in water generally decreases due to the addition of a salt with some exceptions, like glyoxal (Kampf et al., 2013). However, as atmospheric aerosols contain water, salts, and organic compounds, there are likely conditions where the solubility of an organic is more or less favorable in the water–inorganic-rich phase. Mixed organic–inorganic solutions have been observed to phase-separate into an organic-rich and inorganic-rich phase based on their degree of functionalization (as measured by O : C and relative humidity. The O : C serves as a proxy for molar polarization, which dictates the magnitude of the salting-out effect through the Satchenov equation (Bertram et al., 2011). The relative humidity above which a single combined phase exists is called the separation relative humidity. The SRH is not expected to be a strong function of the organic-to-inorganic ratio (Bertram et al., 2011; You et al., 2013), molecular weight of the organic species, or temperature between 244 and 290 K (You and Bertram, 2015). However, the SRH is a function of the type of salt
present, with ammonium sulfate having higher SRH (and less frequent mixing) than ammonium bisulfate, sodium chloride, and ammonium nitrate for a given O : C. During SOAS, inorganic aerosol was dominated by (NH₄)₂SO₄ and NH₄HSO₄, and SRH was diagnosed in CMAQ based on the You et al. (2013) experimental results for ammonium sulfate. The relationship for SRH (fraction between 0 and 1) as a function of O : C was recast in terms of OM / OC:

\[
SRH = \left[1 + \exp\left(\frac{7.7 \text{OM}}{\text{OC}} - 15.8\right)\right]^{-1}. \tag{7}
\]

Since ammonium sulfate has the highest SRH of the salts examined by You et al. (2013), choosing another salt would increase the frequency of phase mixing and difference compared to the base simulation.

For simulations considering phase separation in CMAQ (ideal Wᵢ and γ ≠ 1), when the ambient relative humidity was below the SRH, the model separated the particle into a water-rich phase (containing aqueous SOA) and an organic-rich phase (containing traditional SOA and POA). This separation of aqueous SOA and traditional SOA at low RH is consistent with the work of Ye et al. (2016), who found that isoprene SOA surrogates unfavorably interacted with α-pinene SOA even at 60% RH.

### 2.5 Predicting water uptake to the organic phase

Water uptake to the organic phase (W₀ > 0 and γ ≠ 1 simulations) was predicted in CMAQ using κ-Köhler theory and solving for the volume-equivalent diameter, D (Petters and Kreidenweis, 2007):

\[
\text{RH} - \frac{D^{3} - D_{\text{core}}^{3}}{D^{3} - D_{\text{core}}^{3}(1 - \kappa)} \exp\left(\frac{4\sigma_{w}M_{w}}{RT\rho_{w}D}\right) = 0 \tag{8}
\]

and

\[
D_{\text{core}} = \left(\frac{6}{\pi} \sum_{i \neq W_{0}} V_{i}\right)^{1/3}, \tag{9}
\]

where D_{core} is the volume (V) equivalent accumulation mode diameter excluding water associated with organic species, M_{w} is the molecular weight of water, ρ_{w} is the density of water, R is the universal gas constant, T is temperature, and σ_{w} is the surface tension of water (0.072 J m⁻²). In order to calculate the volume-equivalent diameters, D and D_{core}, particle density was needed. Density values in CMAQ v4.7–v5.1 for organic constituents are generally on the order of 2000 kg m⁻³. The densities of organic aerosol species were updated to chamber-specific information when available (Ng et al., 2007; Chan et al., 2009) and to 1400 kg m⁻³ otherwise. The mass of particle liquid water associated with organic compounds per volume of air (W₀) was calculated from

\[
W_{0} = \frac{\pi N_{p} \rho_{w}}{6} (D^{3} - D_{\text{core}}^{3}). \tag{10}
\]

H. O. T. Pye et al.: Aerosol-water–organic interactions

where N_p was number of particles per volume air. Total aerosol water in the model was computed as the sum of water associated with inorganics (W_i) calculated with ISORROPIA v2.2 (Fountoukis and Nenes, 2007) and W₀.

The hygroscopicity parameter, κ, was calculated as a volume-weighted sum of the individual component κ_i (Petters and Kreidenweis, 2007) ignoring water associated with organics:

\[
\kappa = \frac{\sum_{i \neq W_{0}} (\kappa_{i} V_{i})}{\sum_{i \neq W_{0}} (V_{i})}. \tag{11}
\]

Cloud condensation nuclei (CCN)-based κₛ were used following Lambe et al. (2011) due to the completeness of that study. The O : C values obtained by Lambe et al. (2011) were increased by 27% to account for a low bias in old calibrations (Canagaratna et al., 2015). In addition, the relationship was recast in terms of OM / OC, resulting in

\[
\kappa_{\text{org},i} = 0.11 \frac{\text{OM}}{\text{OC}} - 0.10. \tag{12}
\]

Equations in terms of O : C are available in the Supplement.

For subsaturated conditions, like those relevant to predicting water uptake, the hygroscopic growth factor (hgf) κ is most relevant (Pajunen et al., 2015); however, CMAQ simulations used CCN-based κ_{org,i} to predict water uptake. Hgfs based κₛ from Dullissly et al. (2011) and Raatikainen et al. (2010) were combined with data from Jimenez et al. (2009) into a parameterization by Lambe et al. (2011). After correcting the parameterization to use updated O : C, the parameterization including hgf-based data resulted in one negative κ and three κₛ higher than 0.6 (same as ammonium sulfate), which may be an upper limit on κ_{org,i} (Ervens et al., 2011). Thus, contrary to the typical trend of κ_{CCN} > κ_{hgf}, more than half of the species had κ_{CCN} < κ_{hgf}. Variation from study to study may be higher than κ_{CCN} vs. κ_{hgf} variations, which have been found to be within 30% for many compounds and unable to be resolved using common measurement techniques (Petters and Kreidenweis, 2007).

In the processing of model output, the following equation was used to determine how errors in the concentration of organic compounds ([OA]), κ_{org}, and RH propagated to errors in W₀:

\[
W_{0} = \frac{\rho_{w} [\text{OA}] k_{\text{org}} - 1}{(1/\omega_{w} - 1)}, \tag{13}
\]

with the activity of water (a_w) defined as

\[
a_{w} = \frac{\text{RH}}{\exp\left(\frac{\sigma_{w} M_{w}}{RT \rho_{w} D}\right)}. \tag{14}
\]

### 2.6 Representing the effect of water on semivolatile partitioning

Partitioning of semivolatile organic species into an absorbing medium can be described by a modified Raoult’s law
where \( A_i \) is the aerosol phase concentration of species \( i \) (\( \mu g m^{-3} \) air), \( G_i \) is the gas-phase concentration of \( i \) (\( \mu g m^{-3} \) air), \( M_p \) is the mass of the partitioning medium (\( \mu g \)), \( M_i \) is the molecular weight of the partitioning medium, \( y_i \) is a mole-based activity coefficient, and \( P_{i,\text{sat}} \) is the saturation vapor pressure of pure \( i \). This relationship (Eq. 15) is true regardless of how the partitioning medium is defined. CMAQ, following Schell et al. (2001), defines \( C_i^* \) as

\[
C_i^* = \frac{\tilde{M}_i y_i P_{i,\text{sat}}}{RT},
\]

where the relevant molecular weight is the individual species molecular weight in contrast to the traditional definition of Pankow (1994), which uses the partitioning medium’s molecular weight:

\[
C_i^* = \frac{G_i \tilde{M}_i}{A_i},
\]

where the total moles of the partitioning medium (\( N \)) are

\[
N = N_{\text{other}} + \sum_i (A_i / \tilde{M}_i).
\]

\( N_{\text{other}} \) represents aerosol in the partitioning medium that is not semivolatile during calculation. Including water in the partitioning medium (either from uptake onto hydrophilic organic compounds or from the inorganic phase) increases the moles of partitioning medium by contributing to \( N_{\text{other}} \). The inclusion of water, and even inorganic constituents, in the absorbing phase has been encouraged for simplified models in order to reproduce more detailed calculations (Zueer et al., 2010).

One equation for one unknown can be derived, where \( T_i \) is the total \((G_i + A_i)\) mass of the semivolatile determined by the mass-based stoichiometric coefficients and amount of parent hydrocarbon reacted (\( \alpha \) HC):

\[
f(N) = 0 = \frac{N_{\text{other}}}{N} - 1 + \sum_i \frac{T_i}{C_i^* + M_i N}.
\]

Equation 20 was solved for \( N \) in the model.

\( M_p \approx \tilde{M}_i \) for the interpretation of data from chamber experiments only, and it allows for \( C_i^* \approx C_i^* \) in a single-precursor chamber experiment so that the Odum two-product fit can be determined. Table 2 indicates this was a realistic assumption for most systems as the two surrogate molecular weights vary by less than 10%. This assumption was not necessary within the CMAQ model.

### 2.7 Estimating solubility and deviations from ideality

When deviations from ideality were considered, the saturation concentration used in the modified Raoult’s law was adjusted using an activity coefficient. All organic–organic interactions were assumed to be ideal, and only the inclusion of water drove deviation from ideality. Observations during SOAS indicate that despite a factor-of-7 change in ambient aerosol water concentration from night to day, \( x_w \) (mole fraction of water in the partitioning medium) typically varied over a narrow range (80 to 96% by mole) throughout the day. The activity coefficient for each organic species, \( \gamma_i \), was determined using a one-constant Margules equation:

\[
\ln(\gamma_i) = x_w^2 \ln(\gamma_i^\infty).
\]

Since \( \gamma_i^\infty \) (the temperature-dependent constant in the Margules equation) corresponds to the activity coefficient at infinite dilution in water (\( x_w = 1 \)), it can be estimated based on Henry’s law combined with Raoult’s law:

\[
\gamma_i^\infty = \frac{\tilde{M}_i \rho_w}{H_i C_{0,i}^* RT M_w},
\]

where \( C_{0,i}^* \) is the pure species saturation concentration at \( T \). \( \gamma_i^\infty \) is related to solubility (\( S_i \)) in mass per volume of water:

\[
S_i = H_i C_{0,i}^* RT.
\]

The saturation concentration as a function of water becomes

\[
C_i^* = C_{0,i}^*(\gamma_i^\infty)^{N_w^2/N^2},
\]

where \( N_w \) is the moles of aerosol water in the partitioning medium. This equation applies across the entire organic-to-water spectrum and shows that \( \gamma_i^\infty \) represents \( C_i^* \) of a species in water (\( x_w = 1 \)) normalized to the pure species \( C_{0,i}^* \). Evaluating \( C_i^* \) for pure water provides, \( C_i^* \), the saturation concentration at infinite dilution:

\[
C_i^* = \frac{\tilde{M}_i \rho_w}{H_i RT M_w},
\]

Values are available in Table 2. The solubilities of nonvolatile species derived from traditional precursors (oligomers/accretion products) were estimated based on assuming a \( C_{0,i}^* \) between \( 10^{-2} \) and \( 10^{-5} \mu g m^{-3} \) and the Henry’s law coefficients of Hodzic et al. (2014).

This representation of deviations from ideality resulted in competing effects due to the addition of aerosol water to the partitioning medium. Adding water increased the partitioning medium as described in Sect. 2.6, which led to more SOA. However, adding water also increased the activity coefficient via the Margules model (Fig. S2), leading to higher C and less favorable partitioning (Fig. S3). The Margules model, combined with the fact that all deviations are observed to be positive for the species examined here, indicated...
### Table 2. A priori SOA and SVOC properties: simulation concentration of parent compounds (g) and mass-based stoichiometric yield from parent hydrocarbon reaction (g/cm²).

<table>
<thead>
<tr>
<th>Compound</th>
<th>Mass-based Yield (g/cm²)</th>
<th>SOA Concentration (g/cm²)</th>
<th>SVOC Concentration (g/cm²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>XYL1</td>
<td>1.27</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>XYL2</td>
<td>1.31</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>XYL3</td>
<td>1.36</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>TRP1</td>
<td>1.47</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>TOL1</td>
<td>1.62</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>TOL2</td>
<td>1.65</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>TOL3</td>
<td>1.68</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>PAH1</td>
<td>1.69</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>PAH2</td>
<td>1.70</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>PAH3</td>
<td>1.71</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ORGC</td>
<td>1.72</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>MTNO3</td>
<td>1.73</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>MTHYD</td>
<td>1.74</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ISOPNN</td>
<td>1.75</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>IETET</td>
<td>1.76</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ALK1</td>
<td>1.77</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ALK2</td>
<td>1.78</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**Notes:**
- All a priori SOA and SVOC properties are simulated using the model described in the supplement.
- The mass-based yield (g/cm²) represents the yield of each compound from its parent hydrocarbon reaction.
- SOA and SVOC concentrations (g/cm²) are derived from the model simulations.
that large additions of water reduced SOA due to the activity coefficient adjustment. Indeed, all saturation concentrations for partitioning into pure water \( (C_{a,i}^*) \) are higher than those into pure organic \( (C_{a,i}^t) \) by 1 to 4 orders of magnitude (Table 2). A priori assumptions regarding the solubility and activity of monoterpene nitrates were so nonideal that particulate nitrate was driven entirely out of the particle, inconsistent with observations (Xu et al., 2015a, b). As a result, the Henry’s law coefficient for monoterpene nitrates \( (MTNO_3) \) was increased by a factor of 100 and all activity coefficients were reduced by a factor of 10 compared to a priori values in the CMAQ \( \gamma \neq 1 \) simulation. These adjustments, determined through a series of sensitivity simulations, may have been necessary due to inaccuracies in the Henry’s law coefficients, effects of inorganics, pure species saturation concentrations, molecular weights, Margules model, or a combination of all of the above. A posteriori parameters used in \( \gamma \neq 1 \), which include a factor-of-100 increase in MTNO \(_3\) solubility and factor-of-10 decrease in activity coefficients, are available in Table S6.

### 2.8 Estimating WSOC

WSOC is an operationally defined species measured by adding water to a system and analyzing the dissolved compounds (Sullivan et al., 2004). Particulate compounds with solubilities greater than 10 g L\(^{-1}\) tend to be measured as WSOC regardless of the sampling and extraction method, while compounds with solubilities less than 1 \( \times 10^{-4} \) g L\(^{-1}\) are insoluble (Psichoudaki and Pandis, 2013). To determine the fraction of OA extracted as WSOC \( (WSOC_{p}) \), the particle phase can be modeled as an equilibrium between two phases: \( a \) and \( b \). The fraction of species, \( i \), in phase \( a \) compared to the total particulate species concentration is

\[
f_{a,i} = \left( 1 + \frac{C_{a,i}^s N_b}{C_{b,i}^s N_a} \right)^{-1},
\]

where \( N_a \) and \( N_b \) are the number of moles in phases \( a \) and \( b \), respectively. If phase \( b \) has no water and is ideal, while phase \( a \) is dominated by water and obeys Henry’s law, then the fraction of aerosol species \( i \) extracted as WSOC \( (f_{WSOC,i}) \) is

\[
f_{WSOC,i} = \left( 1 + \gamma_i \frac{WIOA \bar{M}_w}{LWC \bar{M}_i} \right)^{-1},
\]

where WIOA and LWC are concentration of water-insoluble OA and liquid water in mass per volume of air. Thus WSOC depends on the amount of insoluble material, liquid water, Henry’s law coefficient, and pure species saturation concentration.

### 2.9 Observations for evaluation

Simulations were evaluated by comparing to OC from IMPROVE, Chemical Speciation Network (CSN), and South-Eastern Aerosol Research and Characterization (SEARCH) network observations in the eastern US. For comparisons to SEARCH observations, the Jefferson Street, Atlanta, GA (JST), and Birmingham, AL (BHM), urban sites as well as Yorkville, GA (YRK), and CTR, rural sites were considered. In order to estimate secondary organic carbon (SOC), the method of Yu et al. (2007), which uses OC / EC (elemental carbon) ratios, was revised to account for the semivolatile nature of POA. For estimating observed POA from total OA only, POA in CMAQ is assumed to correspond to emissions of \( C_{a,i}^t \approx 3000 \) \( \mu g m^{-3} \) and lower-volatility compounds. The volatility distribution of gasoline vehicle POA from May et al. (2013) and used by the CMAQ-VBS (Koo et al., 2014) was used to estimate how much POA is expected in the particle under ambient conditions.

The fraction of POA in the particle \( (f_P) \) for each observation data point was estimated as

\[
f_p = \sum_{i=1}^5 \frac{\alpha_i}{1 + C_{a,i}^* / (OC_{obs}(OM/OC)_{mod})},
\]

where the volatility profile is described by one nonvolatile and \( C_{a,i}^* \approx 1, 10, 100, \) and \( 1000 \mu g m^{-3} \) surrogate species in the following mass-based abundance \( (\alpha_i) \): 0.27, 0.15, 0.26, 0.16, and 0.17. Observed SOC was estimated from each observed OC by

\[
SOC_{obs} = OC_{obs} - f_P(POC/EC)_{mod}EC_{obs},
\]

therefore,

\[
POC_{obs} = f_P(POC/EC)_{mod}EC_{obs}.
\]

This calculation only accounts for the effect of dilution and partitioning on POC (primary organic carbon) and does not account for chemical processing that may convert POA to SOA. In addition, compared to other volatility profiles such as diesel POA, this profile tends to be weighted toward lower-volatility compounds. As a result, this approach may be an upper bound on the amount of POC (lower bound on SOC).

In addition to the routine monitoring network data, model predictions were compared to data from the CTR (87.25° W, 32.90° N) and Look Rock, TN (LRK; 83.94° W, 35.63° N) sites from the SOAS field campaign in the southeast United States. Observations include water-soluble organic carbon in both particle and gas phase (Xu et al., 2016; Sullivan et al., 2004), aerosol LWC (Nguyen et al., 2014b; Guo et al., 2015), OA (Xu et al., 2015a, b; Budisulistiorini et al., 2015; Hu et al., 2015), and gas-phase species (Nguyen et al., 2015a; Budisulistiorini et al., 2015). The Supplement provides additional evaluation such as a comparison to OH (Feiner et al., 2016), isoprene (Su et al., 2016; Misztal et al., 2017), and 2-methyltetrol (Isaacman et al., 2014; Isaacman-VanWert et al., 2016) concentrations.
3 Results and discussion

3.1 Updated base model

3.1.1 Effect of property updates

Figure 4 shows the updated molecular weights as a function of pure species saturation concentration and colored by OM/OC. Values are summarized in Table 2. Four species that were initially outside the O:C = 0 and O:C = 1 bounds in CMAQ v5.1, ISO1, ISO2, SQT, and BNZ3, were moved within the bounds or just slightly outside as a result of implementing Eq. (2) through (5) for traditional OA. The impact of updated OM/OC and molecular weight had small impacts on OM (up to ~4 % decreases) and larger impacts on OC (5–8 % decrease in OC across the southeast). This change was driven by an increase in the OM/OC of biogenic (semivolatile isoprene and monoterpene) SOA.

Note that there is likely inconsistency in the structure and assumed vapor pressure for 2-methyltetrols and 2-methylglyceric acid. The model considers IEPOX-derived SOA to be mainly 2-methyltetrols and organosulfates with a small amount of oligomers (Pye et al., 2013). All IEPOX-derived species were treated as nonvolatile, but they should be semivolatile given their molecular weight. Lopez-Hilfiker et al. (2016) indicate that IEPOX-derived organosulfates and 2-methyltetrols measured by common techniques include decomposition products of accretion reactions and that IEPOX-SOA should be relatively nonvolatile, consistent with Hu et al. (2016) and the nonvolatile assumption here. The nonvolatile assumption is, however, inconsistent with 2-methyltetrols being present in the gas phase as observed by Xie et al. (2014). The glyoxal SOA in CMAQ also corresponded to a monomeric unit. If oligomers are the dominant form for aqueous methylglyoxal SOA (Altieri et al., 2008), then the molecular weight would need to be increased. Given the nonvolatile nature of IEPOX-derived SOA and glyoxal SOA, they were not significantly affected by the sensitivity simulations.

In the base and updated models, dry deposition of OA played a relatively minor role in removing semivolatile compounds from the system. Volatility was the primary factor determining the relative role of gas vs. particle deposition for a given species with the specific value of the Henry’s law coefficient being less important as indicated by relatively small changes in overall deposition between the base and update (Fig. 5). At 298 K, the less volatile SVOCs became more soluble than predicted by base CMAQv5.1, while the more volatile SVOCs became less soluble. With the new parameters, dry deposition of gas-phase SVOCs increased by 20 %, while wet deposition decreased by 6 %. Total SOA+SVOC deposition changed by less than 2 %, and surface concentrations changed by less than 3 %. Overall, particle-phase deposition accounted for 22 % of the loss of SOA+SVOC mass. Dry deposition of gas-phase SVOCs accounted for 32 %, and wet deposition of gas-phase SVOCs accounted for 46 %. The relative role of gas-phase SVOC wet deposition was twice as important as predicted by Hodzic et al. (2014), consistent with the greater contribution from soluble biogenic species in the southeast in this work. The combined effects of the molecular weight, OM/OC, and deposition updates resulted in a 10 % decrease in predicted OC over the southeast.

3.1.2 Isoprene SOA

Heterogeneously derived IEPOX SOA in CMAQ was assumed to be nonvolatile and, thus, was not greatly affected by the sensitivity simulations. Positive matrix factorization (PMF) analysis of Aerosol Chemical Speciation Monitor (ACSM) data and high-resolution time-of-flight aerosol mass spectrometer (HR-ToF-AMS) data identified a factor with prominent m/z 82 signals (Lin et al., 2012; Xu et al., 2015a, b). This factor was referred to as “IEPOX-OA” and “isoprene-OA”, respectively. While it is largely attributed to IEPOX uptake, it may not be entirely due to IEPOX (Xu et al., 2015a, b; Schwantes et al., 2015). The term “isoprene-OA” will be used to refer to the ambient PMF factor regardless of technique.

Liu et al. (2015) report that only half of the isoprene R O + HO 2 SOA is from IEPOX in laboratory experiments. Furthermore, the AMS isoprene-OA PMF factor is not fully speciated. During SOAS at the CTR site, Lopez-Hilfiker et al. (2016) were able to explain roughly 50 % of the AMS...
isoprene-OA at the molecular level. Hu et al. (2015) explained 78% of isoprene-OA at CTR by molecular tracers measured online (Isaacman et al., 2014) and identified on filters, but only 26% of isoprene-OA was linked to tracers at LRK (Budisulistiorini et al., 2015). The lack of mass closure in these studies may have resulted from a lack of authentic standards for quantifying accretion products (oligomers and organosulfates).

Regional modeling also indicates that a number of later-generation species besides IEPOX contribute significantly to isoprene-derived SOA in the United States. Marais et al. (2016) indicate that isoprene SOA in the eastern US consists mainly of IEPOX (58%) and glyoxal (28%) uptake products, with 14% due to other species. Ying et al. (2015) attribute only 20% of isoprene-OA to IEPOX uptake, with roughly an equal contribution from methylyglyoxal (MGLY) uptake. Semivolatile isoprene SOA and its oligomers accounted for just under 10% of isoprene SOA in their work. Thus, it is unclear if models can consider only SOA from IEPOX for the isoprene system as a surrogate for AMS-measured isoprene-OA.

Figure 6 shows three model definitions of isoprene-OA: SOA due only to IEPOX-reactive uptake, SOA due to IEPOX-reactive uptake and semivolatile isoprene + OH products, and SOA due to IEPOX and glyoxal/merthlyglyoxal uptake. Also included are the PMF factor observations of isoprene-OA from Xu et al. (2015a) for CTR and Budisulistiorini et al. (2015) for LRK. SOA is examined relative to sulfate as sulfate provides the acidity and aerosol medium for heterogeneous uptake (Pye et al., 2013; Marais et al., 2016). Modeled SOA due to IEPOX-reactive uptake was increased relative to CMAQ v5.1 as a result of the higher rate constant for organosulfate formation implemented in this work compared to the work of Pye et al. (2013). At the CTR site, all definitions of isoprene-OA led to overestimates of observed isoprene-OA relative to sulfate. Isoprene-OA based on IEPOX uptake + semivolatile Odum two-product surrogates led to the highest predicted concentrations and a slope of 0.70 compared to the observed slope of 0.45. As a result, CMAQ IEPOX-OA could respond more strongly to changes in SOx emissions than ambient data would suggest as the regression coefficient has been interpreted as the magnitude of the sulfate control on isoprene-OA (Xu et al., 2015a). However, no direct relationship between Odum two-product isoprene SOA and sulfate exists in CMAQ. The correlation between isoprene-OA and sulfate for all three model representations was high ($r > 0.8$) and close to the observed value ($r = 0.91$), which is also consistent with ongoing modeling work with CMAQv5.1 (Vasilakos et al., 2017). On an absolute basis, predicted IEPOX + SV (semivolatile) OA reproduced observed isoprene-OA within 6% overall with small underestimates in the afternoon. However, modeled sulfate concentrations were lower than observed by 30% and ISOPOOH + IEPOX concentrations (Nguyen et al., 2015a) were overestimated by a factor of 2.4 in the model consistent with other modeling work (Vasilakos et al., 2017). Thus, as better agreement is obtained for the gas-phase isoprene species, additional increases in processes that convert isoprene RO2 + HO2 products to the particle phase may be needed despite the overestimates relative to sulfate shown in Figure 6 for Centreville. These additional processes may include accounting for partitioning of 2-methyltetrols to the gas phase, formation of C5-alkene triols, and/or faster oligomerization (Lopez-Hilfiker et al., 2016; Xie et al., 2014; Suratt et al., 2010).

At LRK, the different model representations of isoprene-OA closely resembled the observed isoprene-OA to sulfate ratio of 0.83. IEPOX uptake alone resulted in a slope of 0.61, and IEPOX uptake along with semivolatile isoprene + OH products results in a slope of 0.79. The model showed only a slightly stronger relationship to sulfate than the observations (observed $r = 0.87$), with the different model representations indistinguishable in their correlation with sulfate (model $r = 0.93$ to 0.95). Similar to the model at the CTR site, CMAQ underestimated sulfate at LRK by about 30%. IEPOX + SV isoprene-OA was biased lower than observed isoprene-OA by almost 40%, and the bias in isoprene-OA (IEPOX + SV) was correlated with the bias in sulfate. ISOPOOH + IEPOX was underestimated by 60% at LRK, in contrast to the CTR site, where it was overestimated (Supplement). Thus, isoprene products in the model were too efficiently converted to SOA at LRK despite the low sulfate.

3.1.3 Total OA

Model predictions of OC, SOC, and POC were compared to network observations using the methods described in Sect. 2.9 to determine how model errors in POA (specifically the nonvolatile assumption) could mask errors in...
SOA. An IMPROVE network observation with a value of 16.9 µg C m$^{-3}$ (at SHMM1, Shamrock Mine, CO) had a Cook’s distance (Cook, 1977) much greater than 1 in a base model–observation comparison and was subsequently removed from all analysis. For the IMPROVE network, 86 % of observed OC was predicted to be secondary in nature (Eq. 29), while CMAQ predicted 46 % of OC was secondary. The variability in predicted SOA fraction (standard deviation, $s$, of 0.21) was much higher than the variability in observed SOA fraction ($s = 0.08$). The CSN network (with a greater proportion of urban sites) was slightly less secondary in nature with 79 % of OC as SOC ($s = 0.11$) and CMAQ predicting 40 % of model OC as SOC ($s = 0.19$). The SEARCH network was the most influenced by SOA of the three networks. SEARCH OC was predicted to be 88 % SOC ($s = 0.06$), while CMAQ indicated 58 % SOC ($s = 0.19$). PMF analysis at the urban JST site during summer 2011 and 2013 indicates that POA (hydrocarbon-like organic aerosol (HOA), biomass burning OA (BBOA), cooking organic aerosol (COA)) factors accounted for 18–30 % of total OA (Xu et al., 2015a; Budisulistiorini et al., 2013), while CMAQ predicted a 42 % contribution of POC to OC averaged across the urban and rural sites here.

Figure 7 indicates that overestimates in POC roughly compensated for underestimates in SOC in the updated CMAQ model. CMAQ predicted that total OC was within 20 % of average observed OC across each network. The normalized mean bias (NMB) for POC and SOC was much larger in magnitude than for total OC but relatively constant between networks. Specifically, SOC was lower by 40 %, while POC was higher by a factor of 1.7 to 1.8. The overestimate in model POC at the routine network locations was consistent with the model overestimate in AMS/ACSM-measured POA at SOAS CTR and LRK sites. Neither site resolved a HOA-type aerosol (Xu et al., 2015a; Budisulistiorini et al., 2015), indicating that POA from fossil fuel sources contributed less than 5 % of total OA. A BBOA-type aerosol was resolved at the CTR site and episodic in nature. Comparing CMAQ-predicted POA from all sources to the BBOA factor at CTR indicated that CMAQ generally overestimated POA by a factor of 2, similar to the overestimate for network OC observations.

Additional insight into biases can be obtained by examining the diurnal profiles of OC (Fig. 8). The diurnal profile of observed OC is relatively flat at the SEARCH sites, consistent with flat total OA (Xu et al., 2015b). CMAQ predictions had a pronounced diurnal profile, with higher concentrations (and relatively good performance or overpredictions) at night and lower concentrations (coinciding with underestimates) during the day. Averaged across the two urban sites (JST and BHM), however, CMAQ showed no bias as a result of compensating diurnal and spatial errors. Rural OC (YRK and CTR) was underpredicted by about one-third. Also included in Fig. 8 is the diurnal profile of POC in red dashes. Modeled POC at the Atlanta site correctly showed high concentrations in the morning (06:00 LT) and evening (19:00 LT), but tended to peak several hours earlier than HOA observed at JST in 2012 (Budisulistiorini et al., 2016). JST total model-predicted POC during morning and evening transition hours was roughly the same magnitude as total observed OC, further indicating that CMAQ tends to overestimate primary organic aerosol.

3.2 Role of water

3.2.1 Effect on network OC

Figure 9 shows how including water interactions in absorptive partitioning calculations affected model predictions of OC at routine monitoring network locations. While including water associated with inorganic species (ideal $W_o$ simulation) in the partitioning medium for SOA decreased the bias in SOC for all networks, it led to small increases in the mean error. Except for the SEARCH network, including organic water ($W_o > 0$) also reduced the mean bias at the expense of mean error. The simulation taking into account nonideality ($\gamma \neq 1$) resulted in low normalized mean bias ($\leq 10$ %) and large improvements in the mean bias compared to all other simulations. The mean error for $\gamma \neq 1$ was marginally increased over the base simulation.

Figure 9 highlights that increases in bias occurred during the night (SEARCH network). The largest increases in bias occurred for the $W_o > 0$ simulation as a result of a large contribution of organic water. Similar to the results for the CTR site (Sect. 3.2.3), daytime concentrations of SOC increased but were still low compared to observations. In general, the variability in the bias increased as a result of water interactions, while the mean bias decreased.

Some caution should be applied when comparing model predictions and observations. Measurements of total aerosol mass from IMPROVE and CSN networks are made under relative humidities of 30–50 %, and quartz filters for OC analysis from IMPROVE may be subject to ambient conditions in the field and during shipping before analysis (Solomon et al., 2014). Exposure to low RH could cause evaporation of reversible aqueous SOA (El-Sayed et al., 2016). Kim et al. (2015) have reported that the IMPROVE measurements of OC were 27 % lower than colocated SEARCH measurements during the summer of 2013 and hypothesized the difference to be due to evaporation from the IMPROVE filters during and after sampling. Episodic field campaign observations may be subject to sampling biases as well. Dryers are used ahead of many online aerosol chemistry instruments, and most aerosol water is expected to evaporate in an aerodynamic lens inlet used on many instruments (Zelenyuk et al., 2006; Matthew et al., 2008). Such drying can cause changes in the aerosol phase state (Pajunoja et al., 2016) and could potentially lead to changes in partitioning of soluble organic compounds. El-Sayed et al. (2016) have reported a loss of WSOC after drying. Those authors used a post-drying resi-
dence time of 7 s, which is much longer than the post-drying time used by the AMS in SOAS (approximately 1 s) and the time in the aerodynamic lens (approximately 0.01 s). A prior study reported that evaporation of ammonium nitrate, a water-soluble and semivolatile species, was not observed when using post-drying residence times of about 1 s (Guo et al., 2016). While this topic should be subject to additional research, the AMS data in SOAS are unlikely to have significant biases due to this effect.

### 3.2.2 Frequency of phase separation

Figure 10a shows the June 2013 predicted average OM/OC across the model domain for the simulation in which phase separation and ideal interactions with inorganic water were predicted (ideal $W_i$). Emitted POA in CMAQ has an OM/OC of 1.25 for vehicles, 1.7 for biomass burning, and 1.4 for other sources, and heterogeneous aging of the POA results in the OM/OC increasing with time (Simon and Bhave, 2012). The urban sites of Birmingham, AL, and Atlanta, GA, had predicted OM/OC ratios between 1.3 and 2.2 with a mean of 1.8, while the rural SEARCH sites of Centreville, AL, and Yorkville, GA, had values between 1.7 and 2.2 with a mean of 1.9, consistent with previous work (Simon et al., 2011).

You et al. (2013) found that particles never undergo phase separation for OM/OC above 2.2 (O:C = 0.8) and are always phase-separated when OM/OC is less than 1.8 (O:C = 0.5). Based on Figure 10b and Eq. (7), phase separation was a frequent, but not constant, occurrence. Phase separation was predicted to be more common in urban areas where OM/OC was low and near the western portion of the domain where RH was low. Figure 10c shows RH, SRH, and phase separation for the CTR site. During the day, SRH decreased as a result of increasing OM/OC ratios for both SOA and POA. The increase in frequency of separation during the day was driven by low RH values during the day. At CTR, the highest frequency of phase separation was predicted in the late morning. For other sites, separation was more frequent in the afternoon. These results demonstrate the complexity of aerosol phase behavior in the atmosphere, and this complexity impacts the way observations are collected and interpreted.

![Figure 7](image-url) **Figure 7.** Aerosol OC, POC, and SOC predicted by the base model simulation ($M_5$) compared to CSN, IMPROVE, and SEARCH (JST, BHM, CTR, and YRK) observations ($O_i$). Mean bias ($MB = \sum_{i=1}^{n} (M_i - O_i)$) and mean absolute error (MAE = $\frac{1}{n} \sum_{i=1}^{n} |M_i - O_i|$) are in micrograms of carbon per cubic meter. X symbols indicate mean bias. Boxplots indicate 5th, 25th, median, 75th, and 95th percentile. $r^2$ based on a zero intercept. $n$ is the number of observations. NMB $= \frac{\sum_{i=1}^{n} (M_i - O_i)}{\sum_{i=1}^{n} O_i}$.

![Figure 8](image-url) **Figure 8.** OC as a function of hour of the day for a SEARCH urban (Atlanta, JST) and rural (CTR) site during June 2013. Bars/shading indicate 25th to 75th percentiles. Lines indicate means. Red dashed lines indicate model-predicted POC.

![Graph](image-url)
Model-predicted RH was lower than the observed RH by about 6% (mean bias). Since phase separation occurred when RH was below the SRH, the frequency of separation using model RH was biased high. In addition, since the model used the SRH predicted for ammonium sulfate, predictions further represent an upper bound on the frequency of phase separation. Thus, particles should be internally mixed without phase separation more often than reported in this work. As phase separation was most consistent with default model assumptions, parameterizing the SRH using data from another salt (and using observed RH) would only increase OA as a result of a greater frequency of inorganic water in the partitioning medium.

3.2.3 Effect of water on OA concentrations at CTR

Figure 11 shows the influence of water on aerosols at the Centreville SOAS site during June 2013. The base simulation
underestimated OA overall, but most substantially during the day. Including inorganic water in the partitioning medium when RH > SRH (ideal $W_i$) resulted in increased OA concentrations at all times of day. Reducing phase separation (under ideal conditions in ideal $W_i$ compared to base) has been shown to increase OA concentrations in box modeling (Topping et al., 2013). In CMAQ, concentrations of OA predicted in ideal $W_i$ were 1.5 times higher than observations at night when RH and aerosol liquid water concentrations were highest. Note that nocturnal mixing may be underestimated in the model as indicated by low boundary layer depths, high monoterpene concentrations, and high NOX concentrations compared to observations at night (Pye et al., 2015). The simulation considering uptake of water into the organic phase ($W_o > 0$) produced the highest predicted OA concentrations out of all simulations as a result of feedback in the model. Specifically, uptake of water and inclusion in the partitioning medium caused OA concentrations to increase, which further increased the amount of water in the particle and OA. Daytime OA predictions did not exceed observations, but nighttime model concentrations were a factor of 2 higher than observed. A comparison of model-predicted aerosol water with observed aerosol water (Fig. 11e) indicated that the model overpredicted aerosol LWC by 2–3× at night when interactions were ideal in the $W_o > 0$ simulation.

The simulation accounting for nonideality in addition to phase separation and uptake of water into organic compounds ($\gamma \neq 1$), produced results similar to the simulation considering phase separation and ideal interactions with inorganic water only (ideal $W_i$ in terms of total OA as a function of time of day (model: observation correlation coefficient $= 0.5$; NMB $= 10\%$ ($\gamma \neq 1$), 20% (ideal $W_i$)). However, the composition of the aerosol was different. Both simulations in which water interactions were ideal (ideal $W_i$ and $W_o > 0$) resulted in overpredictions of less oxidized oxygenated aerosol (LO-OOA) and particle-phase organic nitrates (Supplement). Even with the factor-of-100 increase in Henry’s law coefficient for monoterpene nitrates and the factor-of-10 decrease in activity coefficient implemented in $\gamma \neq 1$ compared to a priori estimates, the predicted concentration of organic-nitrate-derived SOA did not substantially change between the base and $\gamma \neq 1$ simulation. The nonideality resulting from including water roughly compensated for the increase in partitioning medium in the case of organic nitrates.

### 3.2.4 Predicting water uptake onto organic compounds

All simulations indicated that OM/OC ratios trended to peak during the day and were near a value of 2, consistent with observations (Fig. 11c). Semivolatile SOA in the model tended to have lower OM/OC ratios than nonvolatile SOA, which resulted in lower OM/OC ratios overall in the sensitivity simulations compared to the base. These differences in OM/OC between the simulations propagated to predicted $\kappa_{org}$ values (Fig. 11d). The base simulation best agreed with the observationally constrained $\kappa_{org}$ values of Cerully et al. (2015), but the model $\kappa_{org}$ was biased low in all simulations.

Basing the $\kappa$ values for organic species on OM/OC (or O:C) may tend to overestimate the $\kappa$ values for organic nitrates (Suda et al., 2014). However, good agreement with the LO-OOA factor ($\kappa = 0.08 \pm 0.02$, Cerully et al., 2015) is obtained for a 50/50 mixture of MTNO$_3$ and its hydrolysis product ($\kappa = 0.09$). The predicted monoterpene SOA $\kappa$ (0.1) is in agreement with laboratory values ($\kappa = 0.03$ to 0.14; Alfarra et al., 2013). In addition, the $\kappa$ for monoterpenes is higher than the $\kappa$ for sesquiterpenes consistent with the trend (but not magnitude) in the work of Alfarra et al. (2013). The $\kappa$ for IEPOX-derived OA (Table 2) was consistent with isoprene-OA value of Cerully et al. (2015) ($\kappa = 0.2 \pm 0.02$) for a 40% organosulfate:60% 2-methyltetrol composition ($\kappa = 0.23$).

Figure 11e shows two observations of aerosol liquid water content compared to model predictions. In the model, aerosol LWC was represented as the sum of water due to inorganic species ($W_i$, referred to as inorganic water) and water due to organic species ($W_o$, referred to as organic water). The ideal $W_i$ and base simulations resulted in the same predictions of aerosol water as only inorganic species were considered in calculating LWC. The difference between the base simulation and observed LWC indicate a potential role for water associated with organic species. The contribution of LWC due to organic species has been estimated as 35% during SOAS with higher contributions (50%) at night (Guo et al., 2015). Both organic and inorganic water were predicted to be highest in concentration during the night or early morning as a result of the diurnal variation in RH.

Both simulations with uptake of water into organic species ($W_o > 0$ and $\gamma \neq 1$) overpredicted LWC at night, with the $W_o > 0$ simulation resulting in greater overprediction as a result of the feedback mentioned earlier. Figure 11f attributes the overprediction in organic water for the $\gamma \neq 1$ simulation to errors in the concentration of OA, hygroscopicity parameter for organic aerosol ($\kappa_{org}$), and $\alpha_w$ (or RH) (Eq. 13). For simplicity in the attribution analysis, RH was converted to activity using a fixed particle diameter of 200 nm (Hu et al., 2016). $W_o$ was not directly measured, but estimated using measured properties. Figure 11f indicates that overestimates in the concentration of OA at night resulted in overestimates in $W_o$. Underestimates in RH and $\kappa_{org}$ decreased the overestimate. Thus, predictions of aerosol water in the sensitivity simulations can be most improved by improving the concentration of OA in the model.

The concentration of organic water and contribution to total aerosol water is shown across the model domain in Fig. 12. $W_o$ was generally predicted to peak in the same locations where OA (Fig. 12c) was high. This trend was not true in locations where RH drove higher or lower water uptake than expected or OA was dominated by fresh POA with low OM/OC. For example, high RH over the Great Lakes.
and off the northeast coast resulted in high concentrations of organic water. High concentrations of OA from fires in Colorado did not translate to high aerosol water as a result of low RH and low OM / OC ratios leading to low \( \kappa_{\text{org}} \). \( \kappa_{\text{org}} \) ranged between 0.11 and 0.14. The contribution to aerosol water resulting from organic vs. inorganic species (Fig. 12b) reflected the ratio of organic-to-sulfate concentrations as aerosol water is proportional to their concentrations.

### 3.2.5 Model relationship to WSOC

The Particle-into-Liquid Sampler (PiLS) instrument used to measure WSOC\(_p\) adds an equivalent volume of water of \( 6 \times 10^6 \mu \text{g m}^{-3}\) air, which is significantly higher than the concentration of aerosol water observed during SOAS at CTR (less than 73 \( \mu \text{g m}^{-3} \); Nguyen et al., 2015b; Guo et al., 2015). Figure 13a shows the fraction of particulate OA present in the aqueous (vs. insoluble) phase (Eq. 27). For the PiLS instrument during SOAS, compounds with \( \gamma_i \approx 100 \) (solubilities as low as 0.1 \( \text{g L}^{-1} \) ) were expected to be part of measured WSOC\(_p\). Biogenic-VOC-derived SOA was particularly soluble, except for potentially MTNO\(_3\). Alkane and aromatic SOA had \( 1000 < \gamma_i \approx 100 \) (0.1 < \( S \approx 10 \)) and, thus, were less soluble. Note that none of the species have very low solubilities, so all SOA species were expected to be at least partially water soluble during extraction depending on ambient conditions. Using the PiLS estimate of the fraction of water-soluble OA of 90\% (Washenfelder et al., 2015), the mole-weighted \( \gamma_i \) for ambient OA was predicted (Eq. 27) to be 100 000 000, much higher than the coefficient predicted for any individual semivolatile constituent in the model.

The base simulation provided a good representation of WSOC\(_p\) at night but underestimated total OC at all hours of the day, particularly during the daytime. \( \gamma \neq 1 \) provided a better estimate of total OC but overpredicted WSOC\(_p\) at night if compounds with \( \gamma_i \approx 1000 \) (\( S > 10 \text{ g L}^{-1} \)) were entirely considered WSOC. Recall that the a priori estimate of solubility for MTNO\(_3\) was increased by a factor of 100 to reconcile modeled and observed LO-OOA and particulate organic nitrate for \( \gamma \neq 1 \). Even with the factor-of-100 increase in Henry’s law, MTNO\(_3\) remained the least soluble biogenically derived SOA species in the model. The large increase in OC for the nonideal simulation was a result of compounds with solubilities greater than 10 \( \text{g L}^{-1} \) or \( \gamma_i \approx 100 \) (Fig. 13c) which were dominated by traditional biogenic SOA and its accretion products. The accretion product from traditional semivolatile SOA is not well constrained in terms of its structure or volatility. In this work, as in the work of Carlton et al.
**Figure 12.** June 2013 mean predicted (a) aerosol water due to organic species, (b) contribution of organic water to total aerosol water, (c) total organic aerosol, and (d) hygroscopicity parameter for the \( \gamma \neq 1 \) simulation.

**Figure 13.** Fraction of OA present in aqueous phase (a) as a function of activity coefficient at infinite dilution and observed OC and WSOC\(_p\) at CTR compared to model predictions (b–c). Panel (a) boxplots indicate observed LWC from Rutgers/NC State (Nguyen et al., 2014b) and Georgia Tech (Guo et al., 2015) during SOAS. For predictions (a), WIOA is 1 µg m\(^{-3}\) and the species molecular weight is set to 180 g mol\(^{-1}\). Predictions in dark grey shading span a factor of 2 in WIOA (0.5 to 2 µg m\(^{-3}\)). Predictions in light grey shading indicate a factor of 10 in WIOA (0.1 to 10 µg m\(^{-3}\)). Panel (b) corresponds to model predictions in the base simulation, while panel (c) corresponds to predictions in the \( \gamma \neq 1 \) simulation. Model predictions of OC are stacked and divided into POC, compounds with \( \gamma_i^\infty > 1000 \) (solubilities less than 10 g L\(^{-1}\), Table 2), monoterpene nitrate OC, compounds with \( \gamma_i^\infty < 1000 \) (solubilities greater than 10 g L\(^{-1}\)), the organic nitrate hydrolysis product, and aqueous SOC.

(2010) and Pankow et al. (2015), the species was assumed to be nonvolatile with an OM/OC of 2 to 2.1. The solubility of low-NO\(_x\) monoterpened-derived species remained above 10 g L\(^{-1}\), even down to species with a saturation concentration of \( 1 \times 10^{-10} \) µg m\(^{-3}\) using Henry’s law coefficient values from Hodzic et al. (2014). If the accretion products (AOLGB) were better represented by a less functionalized species and effectively insoluble, observation–model disagreement in WSOC\(_p\) in \( \gamma \neq 1 \) would be reduced. Another way to reconcile observed and modeled WSOC\(_p\) may
be to take into account deviations from equilibrium during PiLS extraction, which were not considered here.

Figure 14 shows observed water-soluble organic carbon compounds in the gas phase (WSOC$_g$, measured by mist chamber and total carbon analyzer; Hennigan et al., 2009) compared to (a) semivolatile SOA precursors (i.e., those associated with dry organic aerosol in Table 1) and (b) semivolatile and aqueous SOA precursors currently in CMAQ. The figure indicates that considering semivolatile SOA precursors as the only source of WSOC$_g$ in the model underestimated the daytime amount of WSOC$_g$ but that both observed WSOC$_g$ and modeled semivolatile SOA precursors were on the same order of magnitude. Thus, the semivolatile surrogates in the model represented a significant pool of soluble gases. When IEPOX, glyoxal, and methylglyoxal were included in the model estimate of WSOC$_g$, the daytime WSOC$_g$ was slightly overestimated. However, given the factor of 2.4 overestimate in IEPOX+ISOPOOH in the model compared to observations (Supplement), the specification of WSOC$_g$ differed in the model and observations. Figure 14 indicates that during the daytime, either additional water-soluble SOA precursors need to be implemented in the model or the model is correct and a significant portion of ambient WSOC$_g$ does not lead to SOA. Indeed, observedWSOC$_g$ may have large contributions from compounds such as formic acid that are not considered significant SOA constituents (Liu et al., 2012).

Figure 14. Observed WSOC$_g$ (Hennigan et al., 2009; Xu et al., 2016) and model SOA precursors considering only semivolatile surrogates (a) and semivolatile and aqueous surrogates (b). Grey shading represents the interquartile range of the observed data (mean in black). Colors represent different simulations.

4 Conclusions

Current chemical transport models consider the dominant pathways to SOA to be dry processes governed by condensation of low-volatility organic compounds in the absence of water. In addition, models generally do not consider uptake of water by organic species. In this work, the CMAQ model was updated to consider aerosol water interactions with semivolatile SOA species and uptake of water into OA with a focus on simulating conditions during the Southern Oxidant and Aerosol Study of 2013. A method ($\gamma \neq 1$ simulation) was developed to take into account deviations from ideality using an activity coefficient calculated based on the species Henry’s law coefficient, pure species saturation concentration ($C^*_{i\gamma}$), and the mole fraction of water in the particle that resulted in a normalized mean bias of $-4$, $-10$, and $-2\%$ for IMPROVE, CSN, and SEARCH SOC. Monoterpene nitrates were predicted to be the least soluble semivolatile in the model, consistent with SOA yields from $\beta$-pinene + NO$_3$ being comparable under dry and humid conditions (Boyd et al., 2015). However, most biogenic hydrocarbon-derived semivolatile SOA was highly soluble and predicted to be measured as WSOC. Thus, even aerosol formed through dry processes in models may be classified as WSOC as measured by instruments such as the PiLS.

Based on current observations, aerosol water cannot be added to the partitioning medium for semivolatile organic compounds without simultaneously accounting for deviations in ideality. Otherwise, aerosol liquid water and aerosol carbon are overestimated at night. This finding is consistent with the work by Pun (2008), who found that aerosol water concentrations would more than double if ideality was assumed. Hodas et al. (2015) also found that organic–inorganic water-uptake experiments could not be modeled assuming ideal, well-mixed liquids, and assuming ideality overpredicted $\alpha$-pinene SOA concentrations by 100–200% in the work of Zuend and Seinfeld (2012).

All simulations in this work, including the more aggressive ones assuming ideality, could not reproduce daytime observed OA in the southeast US (at SEARCH sites) solely by adding water to the partitioning medium. Including water resulted in increased model error but could reduce the bias in OC. Additional pathways (new precursors and/or new pathways) to OA, particularly during the daytime, are still needed in models.

The updates described here are in three stages of model readiness:

1. Properties of semivolatile OA constituents can immediately be updated in models to be consistent with their assumed volatility and parent hydrocarbon. Base model performance was good in terms of isoprene-OA and total OC compared to routine networks. Property updates in this work (Table 2) are scheduled for public release as part of CMAQv5.2.
2. Prediction of organic water is more uncertain, but OM/OC is a useful proxy and can be used to parameterize water uptake onto organic species via Eq. (12) and $\kappa$-Köhler theory.

3. The effects of water on semivolatile OA partitioning require additional research as deviations from ideality are important. $\gamma_i^{\infty}$ or $C_{ij}^*$ are recommended as useful parameters for characterizing solubility. Models such as the Aerosol Inorganic-Organic Mixtures Functional groups Activity Coefficients (AIOMFAC) model (Zuend et al., 2008) and UManSysProp (Topping et al., 2016) offer opportunities to perform detailed calculations.

In addition, these areas of model improvement are suggested for future work:

1. A treatment of semivolatile primary OA is needed to reproduce observed surrogates for POA. Factor-of-2 overestimates in POA were predicted to compensate for underestimates in SOA on the order of 40% in IMPROVE and CSN networks.

2. Improvements to sulfate and gas-phase isoprene chemistry will lead to an improved isoprene-OA representation in models as isoprene-OA is correlated with sulfate, but precursors to IEPOX-derived SOA were overestimated at CTR during SOAS. Predictions of isoprene SOA could be further improved by considering the volatility of IEPOX-derived species (such as 2-methyltetrols and C$_5$-alkene triols) (Isaacman-VanWertz et al., 2016) as well as formation of additional species (Riedel et al., 2016).

3. Model-predicted aerosol LWC that includes water associated with organic compounds can be most improved by improving the concentration of OA, which may require a number of updates in different areas.

4. New precursors to SOA are likely needed, especially during the day when OA is underestimated and gas-phase semivolatile model species are less plentiful. Additional precursors for the isoprene system may include multifunctional hydroperoxides (Riva et al., 2016).  

5 Code availability

CMAQ is publicly available via github (https://github.com/USEPA/CMAQ/). Property updates will be available as part of CMAQv5.2. The code for the base simulation is available on github (v5.2). The code for sensitivity simulations is available upon request.

6 Data availability

SOAS field data are available at http://esrl.noaa.gov/csd/groups/csd7/measurements/2013senex/. Data in the figures are uploaded as an asset.

The Supplement related to this article is available online at doi:10.5194/acp-17-343-2017-supplement.

Acknowledgements. We thank the Computer Sciences Corporation for emission processing and Shaojie Song for GEOS-Chem simulations. We thank Jesse Bash, Donna Schwede, and Matt Woody for useful discussion and Kirk Baker for developing the SOAS modeling platform. We thank William H. Brune, David O. Miller, Philip A. Feiner, and Li Zhang for providing OH data. We thank Paul Wennberg and Tran Nguyen for providing CTR CIMS data for IEPOX + ISOPOOH. We thank Rohit Mathur and Golam Sarwar for providing comments on the paper. The US EPA through its Office of Research and Development supported the research described here. It has been subjected to Agency administrative review and approved for publication but may not necessarily reflect official Agency policy. The Ng group was supported by NSF grant 1455588 and EPA grant RD-83540301. Rodney Weber was supported by NSF grant 1242258. Weiwei Hu and Jose L. Jimenez were supported by EPRI 10004734, NSF AGS-1360834, and EPA STAR 83587701-0. Annmarie G. Carlton was supported by EPA STAR grant R83512 and NSF-AGS 1242155. The Suratt group was supported by EPRI as well from EPA grant R835404 and NSF Grant CHE-1404644. The Goldstein team at UC Berkeley acknowledge support from EPA STAR Grant R835407 and NSF Grant AGS-1250569. Gabriel Isaacman-VanWertz is supported by the NSF Graduate Research Fellowship (no. DGE 1106400).

Edited by: D. Topping
Reviewed by: two anonymous referees

References


Fountoulakis, K. and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K\textsuperscript{+}, Ca\textsuperscript{2+}-Mg\textsuperscript{2+}-NH\textsubscript{4}\textsuperscript{+}-SO\textsubscript{4}\textsuperscript{2−}-NO\textsubscript{3}\textsuperscript{−}-Cl\textsuperscript{−}-H\textsubscript{2}O aerosols, Atmos. Chem. Phys., 7, 4639–4659, doi:10.5194/acp-7-4639-2007, 2007.


H. O. T. Pye et al.: Aerosol-water–organic interactions