Development of a multiphase chemical mechanism to improve secondary organic aerosol formation in CAABA/MECCA (version 4.5.6-rc.1)

Felix Wieser¹, Rolf Sander², and Domenico Taraborrelli¹

¹Forschungszentrum Jülich GmbH, Institute of Energy and Climate Research, IEK-8: Troposphere, Jülich, Germany
²Atmospheric Chemistry Department, Max Planck Institute for Chemistry, Mainz, Germany

Correspondence: Felix Wieser (f.wieser@fz-juelich.de) and Domenico Taraborrelli (d.taraborrelli@fz-juelich.de)

Abstract. During the last decades, the impact of multiphase chemistry on secondary organic aerosol (SOA) has been demonstrated to be key in explaining lab experiments and field measurements. However, global atmospheric models still show large biases when simulating atmospheric observations of organic aerosols (OA). Major reasons for the model errors are the use of simplified chemistry schemes of gas-phase oxidation of vapors and parameterization of heterogeneous surface reactions. The photochemical oxidation of anthropogenic and biogenic volatile organic compounds (VOC) leads to products that either produce new SOA or are taken up by existing aqueous media like cloud droplets and deliquescent aerosols. After partitioning, aqueous-phase processing results in polyols, organosulfates, and other products with a high molar mass and oxygen content. In this work, we have introduced the formation of new low-volatility organic compounds (LVOC) into the multiphase chemistry box model CAABA/MECCA. Most notable is the addition of the SOA precursors limonene, long-chain alkanes (up to 8 C atoms), and a semi-explicit chemical mechanism for the formation of LVOC from isoprene oxidation in the gas- and aqueous-phase. Moreover, Henry’s law solubility constants and their temperature dependences have been estimated for the partitioning of organic molecules to the aqueous phase. Box model simulations indicate that the new chemical scheme predicts enhanced formation of LVOC, which are accounted for being precursor species to SOA. As expected, the model predicts that LVOC is positively correlated to temperature but negatively correlated to NOₓ levels. However, the aqueous-phase processing of isoprene-epoxydiols (IEPOX) displays a more complex dependence on these two key variables. Semi-quantitative comparison with observations from the SOAS campaign suggests that the model may overestimate methylbutane-1,2,3,4-tetrol (MeBuTETROL) from IEPOX. The extensions in CAABA/MECCA will be ported to the 3D-atmospheric model MESSy for a comprehensive evaluation of the impact of aqueous-phase chemistry on SOA at a global scale.
1 Introduction

Secondary organic aerosol (SOA) is formed from anthropogenic and biogenic VOC in the atmosphere and comprises a significant amount of the total OA mass (Hallquist et al., 2009). Atmospheric aerosols have received increased attention in recent years, as a result of their impact on human health, urban visibility, and climate change (Lin et al., 2017; Chen et al., 2020; Zhong and Jang, 2011; Zhu et al., 2017). New SOA precursors, formation pathways, and loss reactions in the gas and the aqueous phase have been discovered (Wennberg et al., 2018; Hodzic et al., 2014; Lim et al., 2010). In current models, aqueous-phase pathways leading to SOA are continuously improved, as their importance has been demonstrated in experiments (Lin et al., 2014; Carlton et al., 2007; Ervens, 2015). Hu et al. (2015) investigated aqueous SOA (aqSOA) tracers from isoprene oxidation and found a contribution between 6 % and 36 % to total OA.

Model results display different biases in recent models. Traditional global models generally underestimate SOA mass, especially during haze events, due to the increase in pollutant concentrations (Heald et al., 2011; Tilmes et al., 2019). This gap between models and observations is continuously decreasing in recent years (Shrivastava et al., 2017). Aqueous SOA is known to be more oxidized than SOA formed from gas-phase precursors, thus the modeled O/C ratios commonly do not match experimental results (Lim et al., 2010). Additionally, products from aqueous-phase oxidation are formed on different time-scales than from gas-phase oxidation, which impacts the vertical distribution of SOA in the atmosphere (Seinfeld and Pandis, 2016; Hodzic et al., 2016).

We have implemented recent experimental results and formerly neglected SOA precursors into the chemistry scheme of the box model CAABA/MECCA. The updated chemical mechanism is meant to be used for further advancing global simulations of SOA with the EMAC model (Pozzer et al., 2022). EMAC is the global configuration of the Modular Earth Submodel System (MESSy) (Jöckel et al., 2006; Kerkweg et al., 2007; Tost et al., 2006; Jöckel et al., 2010, 2016). This approach has two advantages: 1) model simulations in the test phase are considerably faster in the box than in the global model; 2) thanks to the MESSy interface structure, the MECCA chemical mechanism can directly be used in the global model EMAC. This allows us to be fast in the implementation process while being able to use the new chemistry in the global model without further adaptations. The partitioning to the aqueous/particle phase plays an important role in the SOA formation process. As temperature-dependent Henry’s law solubility constants ($H_s$) for large organic molecules are sparse in the literature, these properties have to be estimated. The corresponding approach is discussed in Sect. 2.3. In the following section, we introduce major specifications of the CAABA/MECCA model.

2 Model description

2.1 Specifications

CAABA/MECCA is a combination of the box model CAABA (Chemistry As A Box model Application) and the atmospheric chemistry model MECCA (Module Efficiently Calculating the Chemistry of the Atmosphere) (Sander et al., 2019). The chemistry module MECCA is also used in the global model EMAC, simplifying the adoption of box model changes into the global
model. MECCA contains both gas- and aqueous-phase chemistry and the species can partition between the phases. For the partitioning, Henry’s law solubility constants are utilized. Available mechanisms for organic chemistry are the Mainz Organic Mechanism (MOM) (Sander et al., 2019) and the Jülich Aqueous-phase Mechanism of Organic Chemistry (JAMOC) by Rosanka et al. (2021). MOM is the default oxidation mechanism of VOC in MECCA. It contains an advanced treatment of chemistry of isoprene (Taraborrelli et al., 2012; Nölscher et al., 2014; Novelli et al., 2020), monoterpenes (Hens et al., 2014; Mallik et al., 2018) and aromatics (Cabrera-Perez et al., 2016; Taraborrelli et al., 2021). JAMOC is a reduced subset of the CLEPS mechanism (Mouchel-Vallon et al., 2017). Although built for global chemistry simulations, both chemical mechanisms together cover a wide range of reactions. Combined, they include gas-phase oxidation and gas/aqueous partitioning for species up to 11 carbons and aqueous-phase oxidation for species up to 4 carbons. The submodel JIVAL calculates photochemical rate constants ($j$ values) based on cross-sections and quantum yields (Sander et al., 2014). The related $j$ values are usually first determined for the smallest compound of a compound class and this value is applied to higher homologs. For instance, $j$ values for the photolysis of all organic hydroperoxides are taken equal to the one for methyl hydroperoxide (CH$_3$OOH). This is also done in the Master Chemical Mechanism (MCM) (Jenkin et al., 1997). More details about CAABA/MECCA can be found elsewhere (Sander et al., 2005, 2011, 2019).

2.2 Mechanism development

2.2.1 Overview

Table 1 lists the newly implemented and updated precursors, split into biogenic and anthropogenic origination. A major update to the chemistry mechanism is introduced by updating the isoprene oxidation scheme. Additionally, the $\beta$-pinene and benzene mechanism are revised. Limonene, IEPOX and long chain alkane mechanisms were newly added. The chemistry schemes of the monoterpenes sabinene, camphene, and carene had been based on $\alpha$-pinene and were not developed for the individual compounds. They are now excluded as they do not fulfill the standard of the implemented chemistry. The impact of these monoterpenes on LVOC is small, as depicted in fig. S9 in the supplement. In the following sections, changes to the specific mechanisms are described in detail.

2.2.2 Gas-phase kinetics

The nitrate radical addition to isoprene is improved based on Vereecken et al. (2021), which has been validated against chamber experiments (Carlsson et al., 2023). Similar products as in the former oxidation mechanism are formed, while product yields are redistributed. Isoprene OH-oxidation under low NO$_x$-conditions is revised by adding the formation of epoxydiols (IEPOX) according to St. Clair et al. (2016) and dihydroxy hydroperoxy epoxides (ISOPBEPX) according to D’Ambro et al. (2017). Both compounds were identified as main SOA precursors (Lopez-Hilfiker et al., 2016; D’Ambro et al., 2017). Furthermore, the gas phase oxidation of IEPOX as described by Bates et al. (2014) is included in the new mechanism. To simulate a wide range of monoterpenes, the oxidation of limonene was added to the model. The update is based on the chemical scheme of the Master Chemical Mechanism (MCM v3.3.1) by Jenkin et al. (1997) (http://mcm.york.ac.uk). If available, reaction rates are
Table 1. List of all newly implemented and updated species, together with the main reactants and main mechanism sources. Compounds are divided into biogenic and anthropogenic origination.

<table>
<thead>
<tr>
<th>Precursor</th>
<th>Implementation type</th>
<th>Main reactant</th>
<th>Mechanism sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biogenic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isoprene update</td>
<td>update</td>
<td>NO$_3$</td>
<td>Vereecken et al. (2021)</td>
</tr>
<tr>
<td>IEPOX new</td>
<td>OH</td>
<td>Petters et al. (2021), Riedel et al. (2016)</td>
<td></td>
</tr>
<tr>
<td>/β-pinene update</td>
<td>OH / ozone</td>
<td>Vereecken and Peeters (2012)</td>
<td></td>
</tr>
<tr>
<td>sabinene / carene /</td>
<td>excluded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>camphene excluded</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Anthropogenic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>benzene update</td>
<td>OH</td>
<td>Xu et al. (2020), Wang et al. (2013)</td>
<td></td>
</tr>
<tr>
<td>pentane, hexane,</td>
<td>new</td>
<td>OH</td>
<td>Sivaramakrishnan and Michael (2009)</td>
</tr>
<tr>
<td>heptane, octane</td>
<td></td>
<td></td>
<td>Atkinson et al. (2008)</td>
</tr>
</tbody>
</table>

re-calculated by structure-activity relationships (SARs), and low-yield pathways are excluded. Furthermore, the mechanism was refined by results of Carslaw (2013) and Vereecken and Peeters (2012). Products important for SOA are highly oxidized large hydroperoxides and ketones. Small adjustments referring to Vereecken and Peeters (2012) were introduced for β-pinene. Anthropogenic SOA precursors are represented by aromatics and large n-alkanes. The oxidation scheme of benzene and toluene is rather detailed, while higher substituted aromatics are treated in a simplified manner (Cabrera-Perez et al., 2016; Taraborrelli et al., 2021). Benzene chemistry is updated according to the results by Xu et al. (2020) who found no evidence for the epoxide channel. We distribute epoxide production into competing pathways (see Fig. S7). This adaptation yields more small oxidized compounds like galoxy. Thus a decrease in SOA mass in absence of cloud processing and efficient oligomer formation is expected. Similar changes for the epoxide channel were found for the OH oxidation of toluene, but at the same time the formation of alternate epoxides was shown (Zaytsev et al., 2019; Wu et al., 2014). MOM treats the oxidation of alkanes up to 4 carbons. Based on the field measurements of McDonald et al. (2018), the mechanism has been extended for long-chain alkanes with sizable emissions (up to n-octane). The n-alkane mechanism is based on the work of Atkinson et al. (2008) and yields hydroperoxides, alkyl nitrates, and organic molecules containing ketone and hydroxyl groups. This mechanism is simplified and thus only covers the oxidation of specific reaction sites, and only one H-abstraction process is considered.

Experimental studies have shown partitioning of multifunctional alkyl nitrates to the aerosol phase (Perring et al., 2013). The yields of alkyl nitrates from the new RO$_2$ + NO reactions for limonene and long-chain alkanes are implemented following the protocol by Sander et al. (2019). These yields depend on the number of heavy atoms, temperature, and pressure (Arey...
et al., 2001; Teng et al., 2015). Thus, the updated model is expected to generate significant amounts of SOA precursors over continental polluted regions during wintertime.

The rate constants for common reactions (e.g., OH-addition and H-abstraction) are taken from predefined functions implemented in MECCA, while rate constants for specific reactions are taken from the literature or are calculated by SAR (Kwok and Atkinson, 1995; Sander et al., 2019). In the isoprene and limonene mechanisms we consider H-shifts in peroxy radicals as described and estimated by Vereecken and Nozière (2020). In this scheme, the H-shift depends on neighboring substituents, yielding mainly highly oxygenated molecules (HOMs). New hydroperoxides either react with OH radicals by H-abstraction to reform the corresponding peroxy radical or decompose into an alkoxy and an OH radical. The decomposition step is mainly applied to molecules with neighbouring reactive groups, e.g., double bonds, which show a high reactivity towards alkoxy radicals, yielding epoxides. This kinetic scheme is key to the IEPOX and ISOPBEPX formation. In the alkane oxidation mechanism alkoxy radicals mainly undergo 1,5-H-shifts. Graphical representations of all new reaction mechanisms are shown in Figs. S1-S8 in the supplement.

2.2.3 Aqueous-phase kinetics

Rate constants and branching ratios in the aqueous phase are taken from Mouchel-Vallon et al. (2017), if available. Similar to the CLEPS 1.0 protocol, H-abstractions by hydroxyl radicals are estimated with the SAR by Monod and Doussin (2008). Only the fastest H-abstraction pathways are considered, even though the SAR provides branching ratios for the different reactive sites. The examination of all CH-bonds would require a general aqueous-phase C-H abstraction scheme in MECCA for all generated products, which is not included in the present model.

Kinetics of the IEPOX reactive uptake (acid catalyzed ring opening) are taken from the supplementary information of Petters et al. (2021), while the branching ratios are extracted from Riedel et al. (2016). In a moderately polluted atmosphere, methylbutane-1,2,3,4-tetrol (MeBuTETROL) is the main product of aqueous IEPOX oxidation, while the corresponding organosulfate is dominant in the presence of sulfate aerosols. As an addition to the production scheme, we developed and implemented a loss mechanism based on Cope et al. (2021), which we similarly applied to ISOPBEPX. Cope et al. (2021) revealed formic acid as the main oxidation product. They proposed a mechanism, explaining the formic acid formation, invoking some reactions to proceed at unusual rates in aqueous media. Figure 1 shows our adapted and revised mechanism. Comparing the typical rate constants of the possible pathways, we redistributed the oxygen abstraction by NO to the HO2 elimination. The latter is nevertheless consistent with the observed formic acid yield. Similar to the mechanism by Cope et al. (2021), Fig. 1 displays the abstraction of the hydrogen atom at C4 in the mechanism. In addition, we have implemented the abstraction of the hydrogen atom at C1 after the HO2 elimination yielding hydroxyacetone (not shown). Oxidation of the latter leads to the formation of the geminal diol of methyl glyoxal. The simplest geminal diol, from hydration of formaldehyde, has been shown to efficiently form formic acid by oxidation in both gas and aqueous phase (Franco et al., 2021). Similarly, the geminal diol of methylglyoxal yields formic acid directly but also indirectly via pyruvic acid. The related chemistry is already available in JAMOC (Rosanka et al., 2021).
Figure 1. Simplified MeBuTETROL aqueous-phase oxidation mechanism by Cope et al. compared to the mechanism applied in MECCA. The mechanism is similarly applied for ISOPBEPX. Aqueous-phase production of formic acid is via oxidation of the geminal diol of methylglyoxal (not shown). The H-abstraction at the red hydrogen (C1; assumed to be 50%) yields additional formic acid.

As an update to earlier work, we have introduced the hydrolysis of organic nitrates in the aqueous phase, formed from the gas-phase oxidation of isoprene with OH after NO addition (Vasquez et al., 2020). Similar to the small adaptation to the isoprene NO$_3$ mechanism, we introduce the outgassing of the oligomers of glyoxal and methyl glyoxal as an update to the JAMOC mechanism. The applied $H_s$ values can be found in Table S2 in the supplement.

2.3 Phase partitioning

2.3.1 Henry’s law solubility constants

Henry’s law describes the partitioning of a compound between the gas and the aqueous phase. In MECCA, Henry’s law solubility constants $H_s$ are defined as

\[ H_s = \lim_{c \to 0} \frac{c}{p} \]

(1)

where $c$ and $p$ are the equilibrium concentration and partial pressure of the compound, respectively. Most values are taken from the compilation by Sander (2015). However, for many species, especially for reaction intermediates, data are not available. Therefore, an estimation method is required. Different approaches are available. The most widely used procedures are Henry-WIN and GROHME. HenryWIN is developed and contributed by the US environmental protection agency (US-EPA, 2012). It
estimates Henry’s law solubility constants based on molecular structure descriptors (bond contribution) (Meylan and Howard, 1991). The training set consists of 345 species with additional 90 chemicals for subsequent regression and results are validated with 72 compounds. The second Henry’s law solubility constant estimation method, called GROHME, is based on a group contribution approach. Similar to HenryWIN, the estimation is executed by multiple linear regression (Raventos-Duran et al., 2010). It uses a training set of 345 species and a validation set of 143 chemicals. Raventos-Duran et al. (2010) compared both methods and concluded that both show higher uncertainties for increasing solubilities. They revealed a large error for multifunctional compounds in HenryWIN using the GROHME validation set. \( H_s \) are overestimated for difunctional molecules, while compounds with more than two functional groups are underestimated in HenryWIN. However, running HenryWIN with the training set of GROHME led to increased precision of the HenryWIN prediction.

Figure 2 displays the estimated Henry’s law solubility constants by GROHME and HenryWIN for the closed-shelled compounds of the newly implemented limonene mechanism. The diagram displays a good agreement between the two methods for a vast majority of the estimations, while a small subset shows a high deviation. 61% of all values lie within an order of magnitude. Larger prediction differences have a smaller impact on compounds with a generally high \( H_s \), as these will partition nearly completely to the aqueous phase in both cases. At low to medium \( H_s \) the compounds which are higher estimated by GROHME have multiple functional groups, containing mainly hydroxyl and carbonyl groups. Compounds predicted to be more soluble by HenryWIN mostly contain nitrate groups, together with further functional groups containing oxygen. A comparison to COSMOtherm values from Wang et al. (2017) does not support either of the estimation methods but lies generally in between the predictions. We use \( H_s \) estimated by GROHME.

Figure 2. Comparison of \( H_s \) estimations by GROHME and HenryWIN for the newly implemented limonene mechanism. Both axis have a logarithmic scale. The black line represents a one-to-one comparison between both methods. The red dotted lines show a deviation of one order of magnitude.
Table 2. B-values of group contribution and correction factors for the estimation of the temperature-dependent Henry’s law solubility constants. All fragments and B-values, except for the hydroperoxy fragment, are taken from Kühne et al. (2005) and converted referring to our Henry’s law solubility constant definition (multiplied by \( \ln(10) \)). The correction factors containing hydroxyl groups are similarly applied for hydroperoxides. The correction fragment name describes the corresponding structure in SMILES notation.

<table>
<thead>
<tr>
<th>Fragment</th>
<th>B-factor / [K]</th>
</tr>
</thead>
<tbody>
<tr>
<td>basis</td>
<td>1202</td>
</tr>
<tr>
<td>Csb / single bond</td>
<td>-60</td>
</tr>
<tr>
<td>Cdb / double bond</td>
<td>541</td>
</tr>
<tr>
<td>H</td>
<td>203</td>
</tr>
<tr>
<td>OH</td>
<td>4145</td>
</tr>
<tr>
<td>O / double bond</td>
<td>2931</td>
</tr>
<tr>
<td>-O-</td>
<td>1966</td>
</tr>
<tr>
<td>ONO2</td>
<td>-811</td>
</tr>
<tr>
<td>OOH</td>
<td>3625</td>
</tr>
</tbody>
</table>

**correction factors**

| C(=O)CO                   | -1538          |
| COCCO                     | -1538          |
| COC(OO)                   | -1538          |
| C(=O)O                    | -3009          |
| C(=O)OO                   | -3009          |

### 2.3.2 Temperature dependence

The van’t Hoff equation describes the temperature dependence of equilibrium constants based on enthalpy. We apply it to Henry’s law and define the temperature-dependence factor \( B \) as

\[
B = \frac{\ln H_s}{\frac{d}{d(1/T)}} = \frac{-\Delta_{sol} H}{R},
\]

where \( \Delta_{sol} H \) is the enthalpy of dissolution, and \( R \) is the gas constant. The substance-specific value of \( B \) is often not known, in particular for large or highly-oxidized compounds. Since it can impact the partitioning considerably, it has to be estimated. Kühne et al. (2005) present an estimation approach for \( B \), based on the linear combination of \( B(X) \) values for predefined molecular fragments, resulting in the final value of \( B \). These fragments can be single atoms (e.g., carbon in the chain) or full functional and larger groups (e.g., hydroxyl groups). Unfortunately, the hydroperoxide moiety has not been included as a molecular fragment by Kühne et al. (2005). As it plays an important role in this study, we have estimated \( B(\text{OOH}) \) as the difference between \( B(\text{ethane}) \) and \( B(\text{ethyl hydroperoxide}) \). The \( B(X) \) factors used in this study are shown in Tab. 2. As an example, Fig. 3 illustrates the calculation of \( B(\text{methylglyoxal}) \).
Figure 3. Example for the estimation of the temperature-dependence factor $B$ for Henry’s law solubility constant of methylglyoxal. The subscript 'sb' denotes a single and 'db' a double bond.

The method assumes that $B(X)$ values for different fragments are additive and that the contribution of a molecular fragment is the same for every molecule it is attached to. To compensate for this, Kühne et al. (2005) introduce correction factors for specific molecular structures that can be applied to the calculated $B$ after estimation. The relevant correction factors for molecules included in the update are listed in Tab. 2. Figure 4 shows estimated values of $B$ compared to literature data. While most molecules are predicted within one order of magnitude, a tendency to overprediction can be observed for some compounds. The outliers are mainly comprised of multi-functional molecules that are affected by correction factors, implying that the corrections are not strong enough to achieve fitting values. Nevertheless, this approach reflects the trend indicated by the data from Sander (2015) and Kühne et al. (2005). This estimation approach could be adjusted using additional and stronger correction factors reflecting more the measured values. Note that $B$ is defined differently in this work and in Kühne et al. (2005). For a direct comparison, a conversion of $B$ is necessary (see Table 2).

Figure 4. Comparison of estimated vs literature temperature dependence factor $B$ of Henry’s law solubility constants. Literature data are extracted from Sander (2015) and Kühne et al. (2005).

Figure 5 shows temperature-dependent Henry’s law solubility constants $H_s$ for alcohols and aldehydes between 270 K and 300 K. Estimated $H_s$ reasonably agree with experimental values in case of the alcohols. The aldehydes display a stronger divergence. This might be due to hydration of the aldehydes in aqueous media, or similar interactions, for which the estimation is not corrected for. Nevertheless, Fig. 5 also illustrates the need for temperature-dependent $H_s$, as the constants vary over one
order of magnitude. This effect is found to be more pronounced for compounds containing multiple functional groups, as more molecular fractions with high $B$ values are applied (see Tab. S2).

Figure 5. Comparison between estimated and experimental $H_s$ as a function of temperature. Solid lines denote experimental and dashed lines estimated values. The black solid line denotes $H_s$ at 298 K with no further temperature correction. Experimental data are taken from Staudinger and Roberts (2001).

3 Results and discussion

3.1 Model parameters/initialization

We have evaluated potential impacts of the new mechanism under varying conditions. Table 4 displays the initial mixing ratios used in all model runs, unless mentioned otherwise. The mixing ratio of isoprene has been fixed to 4 nmol/mol, which is the average measured during the Southern Oxidant and Aerosol Study (SOAS). For the same reason, the average aerosol salt composition during the SOAS campaign is adopted (Xu et al., 2015b). All simulations are executed for a full diurnal cycle. The sensitivity runs are summarized in Tab. 5.

Table 3. Aerosol properties: Chemical composition, liquid water content (LWC) and particle radius. The chemical composition is averaged from Xu et al. (2015b). The LWC is taken from Nguyen et al. (2014)

<table>
<thead>
<tr>
<th>Chemical composition</th>
<th>( \text{NH}_4^+ / \text{µg m}^{-3} )</th>
<th>( \text{NO}_3^- / \text{µg m}^{-3} )</th>
<th>( \text{SO}_4^{2-} / \text{µg m}^{-3} )</th>
<th>LWC / ( \text{µg m}^{-3} )</th>
<th>( r_{\text{aerosol}} / \text{µm} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>0.7</td>
<td>2.4</td>
<td>3.0</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

3.2 BASE vs OLD run

The CAABA/MECCA box model does not simulate the production of SOA particles. Yet, a first assessment of the impact of the newly added chemistry on SOA precursors can be done by analyzing low-volatility organic compounds (LVOC) which
Table 4. Initial mixing ratios for all model runs. Values for NO and NO$_2$ are adapted to low, medium and high emissions in Sect. 3.3. The mixing ratios of O$_2$, N$_2$, CO$_2$ and isoprene are fixed.

<table>
<thead>
<tr>
<th>Species</th>
<th>Mixing Ratio / [nmol/mol]</th>
<th>Species</th>
<th>Mixing Ratio / [nmol/mol]</th>
</tr>
</thead>
<tbody>
<tr>
<td>H$_2$O$_2$</td>
<td>7</td>
<td>MGLYOX</td>
<td>0.5</td>
</tr>
<tr>
<td>O$_3$</td>
<td>25</td>
<td>CSH8</td>
<td>4</td>
</tr>
<tr>
<td>O$_2$</td>
<td>2.1×10$^8$</td>
<td>PAN</td>
<td>0.1</td>
</tr>
<tr>
<td>NH$_3$</td>
<td>1</td>
<td>NO$_3$</td>
<td>3.15×10$^{-3}$</td>
</tr>
<tr>
<td>NO</td>
<td>2×10$^{-2}$</td>
<td>APINENE</td>
<td>0.6</td>
</tr>
<tr>
<td>NO$_2$</td>
<td>4×10$^{-2}$</td>
<td>BPINENE</td>
<td>0.6</td>
</tr>
<tr>
<td>HNO$_3$</td>
<td>5×10$^{-3}$</td>
<td>HCOOH</td>
<td>0.35</td>
</tr>
<tr>
<td>N$_2$</td>
<td>7.8×10$^8$</td>
<td>LIMONENE</td>
<td>0.6</td>
</tr>
<tr>
<td>CH$_4$</td>
<td>1.86×10$^3$</td>
<td>BENZENE</td>
<td>0.1</td>
</tr>
<tr>
<td>HCHO</td>
<td>5</td>
<td>ACETOL</td>
<td>4</td>
</tr>
<tr>
<td>CO</td>
<td>100</td>
<td>CSH12</td>
<td>0.4</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>3.5×10$^5$</td>
<td>C6H14</td>
<td>0.4</td>
</tr>
<tr>
<td>CH$_3$CO$_2$H</td>
<td>2</td>
<td>C7H16</td>
<td>0.4</td>
</tr>
<tr>
<td>CH$_3$CO$_3$H</td>
<td>1.5</td>
<td>C8H18</td>
<td>0.4</td>
</tr>
<tr>
<td>CH$_3$OH</td>
<td>0.5</td>
<td>CH$_3$OOH</td>
<td>4</td>
</tr>
<tr>
<td>HONO</td>
<td>4×10$^{-5}$</td>
<td>TOLUENE</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 5. Abbreviation and description of the different setups for the sensitivity runs.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASE</td>
<td>base run with the updated mechanism</td>
</tr>
<tr>
<td>OLD</td>
<td>model run with CAABA/MECCA 4.5.5</td>
</tr>
<tr>
<td>BASE-278K</td>
<td>low temperature run with the updated mechanism</td>
</tr>
<tr>
<td>OLD-278K</td>
<td>low temperature run with CAABA/MECCA 4.5.5</td>
</tr>
<tr>
<td>High-NO$_x$</td>
<td>model run with the updated mechanism at high NO$_x$</td>
</tr>
<tr>
<td>Medium-NO$_x$</td>
<td>model run with the updated mechanism at medium NO$_x$</td>
</tr>
<tr>
<td>Low-NO$_x$</td>
<td>model run with the updated mechanism at low NO$_x$</td>
</tr>
</tbody>
</table>

can act as SOA precursors. Thus, we analyze the total (gas, aqueous and aerosol) mixing ratios of implemented LVOC. In the scope of this analysis, LVOC are defined as VOC fulfilling the condition $H_s > 10^8$ M/atm, following Hodzic et al. (2014). For evaluating the impact of the new mechanism on VOC in different molecular size ranges, the total LVOC has been subdivided into three groups: small-LVOC (up to four carbons), medium-LVOC (five and six carbons) and large-LVOC (more than six carbons). The new mechanism has a marginal impact on the simulated small-LVOC. Hence, this class is not included in the following discussion. Figure 6 shows mixing ratios of the BASE and the OLD sensitivity run at 298 K and 278 K, respectively.
Mixing ratios of key radicals during all sensitivity runs are depicted in the supplement (see Fig. S10-S16). The medium-LVOC displays the largest impact. Before sunrise (prior 8 h), no substantial mixing ratios are depicted. After sunrise, mixing ratios increase for the BASE and the OLD run, while the BASE run displays an approximately eight times higher peak mixing ratio. This is mainly due to the formation of MeBuTETROL and organosulfates from isoprene as shown in Fig. 7. The temperature change to 278 K lowers the medium-LVOC yield for both sensitivity runs, while the ratio between both remains similar. For the BASE run, the concentration decrease with temperature can be explained by a decrease in IEPOX formation in the gas phase, due to slower gas phase oxidation rates (see OH mixing ratios in Fig. S10-S11), with a simultaneous decrease of the acidity in the aqueous phase (see Fig. S11). With the main SOA pathway of IEPOX being the acid-catalyzed ring-opening, a lower LVOC yield is expected. Nevertheless, the ratio of IEPOX between aqueous and gas phase increases, due to a higher partitioning coefficient. In the global model, an increased production of SOA precursors from isoprene oxidation is expected (see Carlton et al. (2009)).

Large-LVOC also increase during the full simulation period with the introduced update, however, a less pronounced change is predicted. Even though limonene and its oxidation pathways are newly introduced and show non-negligible LVOC yields, the
exclusion of camphene, sabinene, and carene compensates partly for the additional LVOC. For all three, a mechanism similar to α-pinene had been assumed, due to structural similarities. The oxidation of these monoterpenes will be re-introduced as soon as new experimental/theoretical results are accessible, including a compelling mechanism. Only before sunrise, when oxidation by ozone and NO3 is dominant, the OLD run out-competes the BASE run. This indicates a low LVOC yield in the limonene ozone mechanism. The total LVOC show similar results as the medium-LVOC, together with a positive offset of both runs due to small-LVOC. From this analysis a significant increase in SOA yield can be expected when the mechanism will be incorporated in the global model. Considering the reacted VOC and the produced LVOC in the BASE run we can calculate a LVOC yield to approximate the SOA yield. This analysis results in a yield of 7.3%. With fixed isoprene concentration, it is the main LVOC contributor (approx. 90%). Carlton et al. (2009) collected data from various studies investigating the SOA yield from isoprene and found yields up to 6%. Taking into account that some loss processes are not implemented yet and that the aerosol formation process is not modeled, the agreement between model and experiment is reasonable.

### 3.3 NO\textsubscript{x} dependence

To estimate how anthropogenic NO\textsubscript{x} emissions impact LVOC (and SOA) formation, we evaluate LVOC mixing ratios under varying NO\textsubscript{x} concentrations. The main reactions of NO with peroxy radicals are O-abstraction and addition, forming alkoxy radicals and nitrates, respectively. These processes compete with the formation of hydroperoxides involving HO\textsubscript{2}. As the products from NO oxidation are generally more volatile compared to hydroperoxides, less LVOC are formed under high NO conditions. This hypothesis is supported by Pye et al. (2010), who estimated a low contribution of "RO\textsubscript{2} + NO"-reactions to the total aerosol fraction, while the "RO\textsubscript{2} + HO\textsubscript{2}"-reaction shows a higher fraction for all considered compounds (biogenic VOC). However, this does not always hold true for aromatic compounds. Based on chamber studies, Xu et al. (2015a) found that the SOA yield and NO\textsubscript{x} levels were correlated for toluene but anti-correlated for m-xylene. Figure 8 displays the model results of the sensitivity runs for the low-, medium-, and high-NO\textsubscript{x} scenarios. NO initial mixing ratios were set to 20 pmol/mol for low, 2 nmol/mol for medium and 20 nmol/mol for modeling highly polluted areas. NO\textsubscript{2} mixing ratios are fixed to double the amount of NO mixing ratios. The large-LVOC display the expected trend. With decreasing NO\textsubscript{x} concentration, the mixing
ratios increase as a result of the higher hydroperoxide product share. Especially large-LVOC formation before sunrise from ozone is strongly \(\text{NO}_x\) dependent. The medium-LVOC show a more complex behavior. LVOC mixing ratios rise between low and medium \(\text{NO}_x\) concentrations but fall off for high \(\text{NO}_x\). This can be explained by the relative change of the OH concentration and the hydroperoxide yield. The OH concentration rises with increasing \(\text{NO}_x\), while the hydroperoxide yield continuously decreases. IEPOX and many other LVOCs require the formation of an intermediate or product hydroperoxide group. At medium \(\text{NO}_x\) levels, the rise in OH concentration overshadows the decreasing hydroperoxide yield for the dominant LVOC species. This trend reverses at high \(\text{NO}_x\) concentrations. Due to the comparably high total mixing ratios of the medium-LVOC bin, the total LVOC results reflect mainly the medium-LVOC.

### 3.4 Aqueous-phase chemistry

In the evaluation of the influence of temperature and \(\text{NO}_x\) on LVOC, the aqueous oxidation of IEPOX displays a key role and increases the complexity of the reaction system. Although the box model neglects many environmental effects and dependencies, we want to compare modeled with measured mixing ratios to see whether the model produces tracer compounds in
realistic amounts. For that reason, the model setup was adjusted to the conditions during the SOAS campaign. In this campaign, Hu et al. (2015) reported measurements of MeBuTETROL together with total IEPOX-SOA concentrations (combined concentrations of MeBuTETROL, C5-triols and isoprene-derived organic sulfates) in ambient aerosol. Assuming that the main IEPOX-SOA contributors are MeBuTETROL and the corresponding organosulfates (OS), the OS concentrations can be derived by subtracting the MeBuTETROL from the total IEPOX-SOA concentration. We extracted modeled concentrations after a full model day to compare the results. Note that this comparison is only intended to show whether the model yields results of the same order of magnitude. In measured mass concentrations, only MeBuTETROL found in aerosols is taken into account and general loss pathways are accessible (e.g. deposition). The model, on the other hand, adds up all products formed (gas and aqueous phase), and only the chemical loss of MeBuTETROL is considered. Wet and dry deposition is neglected in the model runs. Further, the modeled OH concentration exceeds the measurements (1.2E-13 vs 5.5E-14 mol/mol, see Fig. S17 and Sanchez et al. (2018)), while the pH is higher in the model (Fig. S11) than in the SOAS aerosol (Guo et al. (2015), median pH of 1). The acid catalyzed ring-opening involving NO$_3^-$ as nucleophile, described in Eddingsaas et al. (2010), is also not considered. Thus, an over-prediction by the model is expected. Table 6 displays the mean measured and modeled MeBuTETROL, OS, and total IEPOX-SOA concentrations, with and without aqueous-phase degradation of MeBuTETROL. Modeled mass concentrations exceed the measured concentrations by an order of magnitude, while the MeBuTETROL over-prediction is 50% higher without aqueous degradation. Considering the influence factors discussed above, this assessment lends confidence that the model for IEPOX-SOA is realistic. But it also stresses the importance of aqueous-phase degradation pathways in models. Non-negligible loss processes are still missing in the MECCA chemical scheme. The aqueous processing of all hydroperoxides, alkyl nitrates, and -sulfates is not implemented, resulting in an overestimation of those species. Additionally, there is no SAR method dealing with the reactivity of epoxides towards OH radicals in aqueous media. Reaction rates have to be estimated by treating epoxides like hydroxyl groups, not considering the enthalpy gained by prompt ring opening. On the other hand, organic coatings on aqueous aerosols and the aerosol phase state may limit the reactive uptake of IEPOX and thus the production of MeBuTETROL and OS (Zhang et al., 2018; Octaviani et al., 2021).

Table 6. Comparison between measured and modeled MeBuTETROL, OS and total IEPOX-SOA mass concentrations. Measured values are taken from Hu et al. (2015). The concentrations were investigated over multiple days, but a mean concentration of 0.3 $\mu$g m$^{-3}$ is assumed in the scope of this comparison. Modeled mixing ratios are extracted after a full modeled day.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Measured / $\mu$g m$^{-3}$</th>
<th>Modeled / $\mu$g m$^{-3}$</th>
<th>Modeled no MeBuTETROL degrad. / $\mu$g m$^{-3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hu et al. (2015)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MeBuTETROL</td>
<td>0.3</td>
<td>4.4</td>
<td>6.6</td>
</tr>
<tr>
<td>OS</td>
<td>0.6</td>
<td>5.3</td>
<td>4.4</td>
</tr>
<tr>
<td>total IEPOX-SOA</td>
<td>0.9</td>
<td>9.7</td>
<td>11.0</td>
</tr>
</tbody>
</table>

In addition to the LVOC concentrations, a change in the O/C ratio might occur as a result of the chemistry update. Especially, because LVOC from aqueous oxidation is added, which commonly possesses a higher O/C ratio (Lim et al., 2010). The main LVOC components, MeBuTETROL and OS, exhibit a ratio of 1.0 and 1.6, respectively. Thus, an increase in the O/C ratio
is expected. Figure 9 displays the O/C ratio of the summed-up LVOC for the BASE and OLD run over the simulation time. After a full modeled day, an increase of 18% is found, with a ratio of 0.76 and 0.93, for BASE and OLD, respectively. The updated model is in good agreement with the SOAS campaign average O/C ratio of 0.91 (Massoli et al., 2018). More accurate degradation pathways of MeBuTETROL and OS are expected to further improve the O/C ratio.

In addition to the already described aqSOA formation sources, there are further known, but missing processes not covered by the mechanism. Formation of nitroaromatics from the aqueous-phase oxidation of anthropogenic VOC is an important additional source of aqSOA. For instance, it has been shown that 4% to 20% of total SOA in Europe is aqSOA originating from residential wood burning (Gilardoni et al., 2016). Accounting for this source of aqSOA requires the development of an explicit mechanism for heterogeneous reactions of $\text{N}_2\text{O}_5$, where chloride and phenols compete with water for the addition to $\text{NO}_2^+$ (Heal et al., 2007; Ryder et al., 2015; Hoffmann et al., 2018; Staudt et al., 2019). A simplified scheme for the nitration of phenol has recently been implemented into MECCA by Soni et al. (2023). Moreover, consideration of nitroaromatics during nighttime oxidation of furans from biomass burning will further increase the model predictions of SOA mass (Joo et al., 2019; Al Ali et al., 2022).

4 Summary and Outlook

The chemical scheme of MECCA has been updated to improve the production of SOA precursors in the model. In addition, we investigated the gas/aqueous phase partitioning of the new species. To assess the update, we investigated the production of LVOC at different physical conditions and initial concentrations in a box model. As expected, we find an increase in the total mixing ratio of LVOC, while the change in LVOC of different sizes is more pronounced. Results display a rising production of medium- and large-sized LVOC, while the number of small LVOC stays constant. This is due to the new implementation of exclusively medium to large precursors. By changing the temperature from 298 K to 278 K, the model predicts a decrease in LVOC production, while an increase in SOA yield is expected. The $\text{NO}_x$-dependence shows more complex patterns of change. In general, the LVOC yield decreases with rising $\text{NO}_x$, but for medium-LVOC, an increase at moderate $\text{NO}_x$ is found. This is due to the aqueous-phase IEPOX oxidation scheme. The modeled O/C ratio confirms the experimental results of the SOAS...
campaign. Overall, we find that the model responds differently to newly added aqueous- and gas-phase reactions. To better reflect the processing of VOC and LVOC in clouds and aerosols, more aqueous-phase reactions, also for the small and large compounds are needed. Furthermore, the implementation of recent findings concerning the peroxy radical reactivity might lead to additional reaction pathways or altered branching ratios. H-shifts of peroxy radicals are only considered in the new limonene mechanism and might lead to an increased formation of HOMs in the model if included for all compounds (Wu et al., 2021).

Similarly, the treatment of peroxy radical dimerization was found to be more important than previously estimated (Schervish and Donahue, 2020). In future work, we want to evaluate the impact of the new mechanism on SOA precursors with explicit multiphase kinetics in deliquescent aerosols and cloud droplets in the global model EMAC. To investigate the dependence of results on the partitioning scheme and the used Henry’s law solubility constants at the global scale, sensitivity runs will be executed with varying partitioning scheme. Further, the introduction of more SOA/LVOC loss processes to the model has shown to be important, as already pointed out by Hodzic et al. (2016). A general aqueous phase degradation scheme of organic nitrates and hydroperoxides would further refine SOA processes.

Code availability. The updated MECCA model code is available as a community model published under the GNU General Public License (https://www.gnu.org/copyleft/gpl.html). The model code can be found in the Supplement (DOI:10.5281/zenodo.7944174) and in the code repository at https://gitlab.com/RolfSander/caaba-mecca. In addition to the complete code, a list of chemical reactions including rate constants and references (meccanism.pdf) and a user manual (caaba_mecca_manual.pdf) are available in the manual directory of the supplement. A list of all Henry’s law and accommodation constants (chemprop.pdf) is available in the tools/chemprop directory. For further information and updates, the CAABA/MECCA web page at http://www.mecca.messy-interface.org can be consulted.

Author contributions. FW and DT designed the study and developed the chemical mechanism. The latter was reviewed by all co-authors and implemented into MECCA by FW and RS. FW performed the estimation of the Henry’s law coefficients. The manuscript was prepared by FW and reviewed by all co-authors.

Competing interests. Some authors are members of the editorial board of GMD. The peer-review process was guided by an independent editor, and the authors have also no other competing interests to declare.

Acknowledgements. The authors gratefully acknowledge the computing time granted through JARA on the supercomputer JURECA (Jülich Supercomputing Centre, 2021) at Forschungszentrum Jülich.


