Atmospheric
Chemistry
and Physics

Testing secondary organic aerosol models using smog chamber data for complex precursor mixtures: influence of precursor volatility and molecular structure

S. H. Jathar[1,*]**, N. M. Donahue**[1]**, P. J. Adams**[1]**, and A. L. Robinson**[1]

[1]Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA, USA
[*]now at: Civil and Environmental Engineering, University of California, Davis, CA, USA

Correspondence to: A. L. Robinson (alr@andrew.cmu.edu)

Abstract. We use secondary organic aerosol (SOA) production data from an ensemble of unburned fuels measured in a smog chamber to test various SOA formation models. The evaluation considered data from 11 different fuels including gasoline, multiple diesels, and various jet fuels. The fuels are complex mixtures of species; they span a wide range of volatility and molecular structure to provide a challenging test for the SOA models. We evaluated three different versions of the SOA model used in the Community Multiscale Air Quality (CMAQ) model. The simplest and most widely used version of that model only accounts for the volatile species (species with less than or equal to 12 carbons) in the fuels. It had very little skill in predicting the observed SOA formation ($R^2 = 0.04$, fractional error = 108 %). Incorporating all of the lower-volatility fuel species (species with more than 12 carbons) into the standard CMAQ SOA model did not improve model performance significantly. Both versions of the CMAQ SOA model over-predicted SOA formation from a synthetic jet fuel and under-predicted SOA formation from diesels because of an overly simplistic representation of the SOA formation from alkanes that did not account for the effects of molecular size and structure. An extended version of the CMAQ SOA model that accounted for all organics and the influence of molecular size and structure of alkanes reproduced the experimental data. This underscores the importance of accounting for all low-volatility organics and information on alkane molecular size and structure in SOA models. We also investigated fitting an SOA model based solely on the volatility of the precursor mixture to the experimental data. This model could describe the observed SOA formation with relatively few free parameters, demonstrat-

ing the importance of precursor volatility for SOA formation. The exceptions were exotic fuels such as synthetic jet fuel that expose the central assumption of the volatility-dependent model that most emissions consist of complex mixtures with similar distribution of molecular classes. Despite its shortcomings, SOA formation as a function of volatility may be sufficient for modeling SOA formation in chemical transport models.

1 Introduction

Secondary organic aerosol (SOA) is aerosol mass formed from the oxidation of gas-phase organic species emitted by natural and anthropogenic sources. Most SOA models underpredict SOA formation both in the laboratory and in the atmosphere (Volkamer et al., 2006; Robinson et al., 2007; Carlton et al., 2010). This indicates that there are large gaps in understanding the numerous precursors and pathways to SOA formation (gas-phase oxidation, multi-generational aging, heterogeneous chemistry, condensed-phase reactions, cloud processing).

Gas-phase oxidation of organic compounds is a major source of SOA, and hence has been extensively studied using smog chamber experiments. However, these experiments have focused on only a small subset (tens) of the thousands of organic compounds found in the atmosphere (Goldstein and Galbally, 2007). SOA models based on those data mostly include emissions and subsequent SOA formation from high-flux volatile organic compounds (VOC) like isoprene, terpenes, single-ring aromatics, and alkanes and alkenes that

have less than or equal to 12 carbons. Odum et al. (1997) were able to explain the SOA formation from unburned gasoline by using a similar model that accounted for SOA formation from the oxidation of single-ring aromatics in the fuel. However, multiple studies have shown that these SOA models are able to explain only a fraction of the SOA measured from diesel and aircraft exhaust, wood smoke and urban plumes (De Gouw et al., 2005; Volkamer et al., 2006; Robinson et al., 2007; Grieshop et al., 2009; Jathar et al., 2012). One contributor to this discrepancy is that combustion sources emit substantial amounts of lower-volatility organics (e.g., hydrocarbons with more than 12 carbons), which are difficult to speciate and therefore not commonly included in models (Robinson et al., 2007). Hence, more research is needed to test these SOA models against data for complex mixtures that contain both small and large organics.

Single-compound smog chamber studies have demonstrated that SOA formation depends both on precursor volatility (vapor pressure) and molecular structure. In general, SOA formation increases with decreasing volatility but has a complex dependence on molecular structure. For example, SOA formation from alkanes increases as the carbon number increases (Lim and Ziemann, 2009; Presto et al., 2010; Tkacik et al., 2012); increasing the carbon number reduces the volatility of the precursor. The importance of molecular structure is illustrated by experiments with different classes of alkanes. For the same carbon number or volatility, cyclic alkanes form the most SOA, followed by n-alkanes, followed by branched alkanes (Lim and Ziemann, 2009; Tkacik et al., 2012). Single-ring aromatics (benzene, toluene and xylenes) form much more SOA than similar sized (C_6 to C_9) n-alkanes or alkenes (Forstner et al., 1997; Ng et al., 2007; Song et al., 2007; Hildebrandt et al., 2009; Lim and Ziemann, 2009). However, multi-ring aromatics and n-alkanes with similar carbon numbers form similar amounts of SOA (Chan et al., 2009; Shakya and Griffin, 2010; Presto et al., 2010).

Existing gas-phase chemical mechanisms account for some differences in volatility and molecular structure. For example, the SAPRC gas-phase mechanism (Carter, 2007) has different model species to account for differences in volatility, e.g., ALK4 (C_5–C_7) versus ALK5 (C_{7+}), and molecular structure, e.g., alkanes (ALK) versus alkenes (OLE) versus aromatics (ARO), so SOA models built using SAPRC can account for some differences in volatility and molecular structure. However, SAPRC and other reduced chemical mechanisms were developed to simulate ozone formation and therefore lump different species by reaction rates rather than SOA mass yields. Even fully explicit mechanisms such as the Master Chemical Mechanism (Saunders et al., 2003) emphasize low carbon-number VOCs that dominate the hydrocarbon flux in most locations, due to a focus on ozone formation and fast photochemistry. Therefore, all of these mechanisms have limitations for predicting SOA formation from complex mixtures.

Another challenge is incorporating unspeciated organics into SOA models. Low-volatility organics (semi-volatile and intermediate-volatility organic compounds; SVOC and IVOC) are thought to be important classes of SOA precursors (Robinson et al., 2007) but are difficult to speciate (Schauer et al., 1999, 2001, 2002; Rogge et al., 1993, 1998; Fraser et al., 1997). Gas-particle partitioning (Robinson et al., 2007; May et al., 2013b, c; May et al., 2013a) and gas-chromatography (Presto et al., 2011) measurements provide some information on the volatility distribution of these emissions; therefore, chemical transport models (CTMs) have adopted a volatility-based approach to model SOA formation from unspeciated S/IVOCs. For example, Robinson et al. (2007) assumed that unspeciated SVOCs and IVOCs and their products react with the hydroxyl radical (OH) to form products that were one order of magnitude lower in volatility than their precursor. Jathar et al. (2012) extended this approach by distributing the reaction products across a range of volatilities. Pye and Seinfeld (2010) proposed a single-step mechanism for SVOC where the products of oxidation were two orders of magnitude lower in volatility than the precursor and used SOA mass-yield data for naphthalene as a surrogate for unspeciated IVOCs. Although these studies have adopted somewhat different approaches, they all show that including unspeciated organics in SOA models helps close large gaps between predicted and measured SOA mass concentrations (Shrivastava et al., 2008; Tsimpidi et al., 2009; Dzepina et al., 2010; Pye and Seinfeld, 2010; Jathar et al., 2011). These new models used a lump-and-yield scheme to model SOA formation from speciated emissions and a volatility-based scheme to model SOA formation from unspeciated S/IVOCs. Given the success of volatility-based schemes, it may be worthwhile to explore their capabilities for modeling all of SOA, not just SOA from S/IVOCs.

In this work, we compared predictions from two different types of SOA models to published data from smog-chamber experiments conducted with different types of unburned fuel (Jathar et al., 2013): (1) three different variants of the SOA model that used volatility- and molecular structure-resolved schemes used in the Community Multiscale Air Quality (CMAQ) model and other chemical transport models and (2) a volatility-dependent model that relates SOA production only to the precursor volatility and ignores molecular structure.

2 Materials and methods

2.1 Experimental data

Jathar et al. (2013) present data from twenty-three high-NO_x photo-oxidation experiments conducted in the Carnegie Mellon University smog chamber to quantify the SOA formation from eleven different fuels (gasoline: 3 experiments; Fischer–Tropsch derived from coal: 2 experiments;
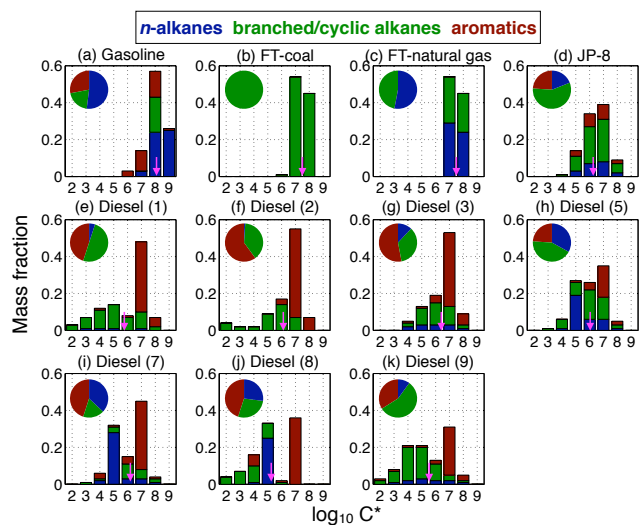
Figure 1. Volatility and molecular structure distributions for unburned **(a)** gasoline, **(b)** Fischer–Tropsch from coal, **(c)** Fischer–Tropsch from natural gas, **(d)** Jet Propellent-8 and **(e)**–**(k)** Diesel 1, 2, 3, 5, 7, 8 and 9, represented in the volatility basis set. For each plot, the bars sum up to 1. The inset pie shows the relative fractions of *n*-alkanes, branched/cyclic alkanes and aromatics in the fuel. The magenta arrow shows the mass-weighted average of the volatility distribution.

Fischer–Tropsch derived from natural gas: 2 experiments; Jet Propellent-8: 6 experiments; multiple diesels: 10 experiments). The Fischer–Tropsch fuels are synthetic substitutes for Jet Propellent-8 fuel. The diesels are part of a test fuel matrix (Fuels for Advanced Combustion Engines) designed to investigate limits in possible fuel composition around three properties of importance to combustion engines: ignition quality, fuel chemistry and volatility (Alnajjar et al., 2010).

The fuels span a wide range of volatility and molecular structure. In Fig. 1 we plot the fuel composition data using the volatility basis set (VBS) (Donahue et al., 2006), color-coded by molecular structure. The VBS is a modeling framework that classifies organics into logarithmically spaced bins of effective saturation concentration (C^*). C^* (inverse of the partitioning coeffcient, K_p (Pankow, 1994)) is proportional to the saturation vapor pressure; it is a semi-empirical property describing the gas-particle partitioning of an organic mixture at 298 K.

To represent a fuel in the VBS, one needs to know the volatility (or C^*) of each individual species in the fuel. However, the granularity of the composition data varied across the fuels. The most comprehensive data were available for the diesel fuels (Alnajjar et al., 2010), which had finely resolved data both by organic class and carbon number. For gasoline, we had alkane and aromatic data by carbon number but no information about the specific species and no resolution in the alkene data. For Fischer–Tropsch from natural gas and Jet Propellent-8, we had speciated data for *n*-

alkanes ranging from *n*-heptane to *n*-nonadecane but lumped data for branched alkanes, cyclic alkanes and aromatics. We assumed that the unspeciated hydrocarbons in the Fischer–Tropsch from natural gas and Jet Propellent-8 have the same carbon number distribution as the *n*-alkanes in that fuel. Finally, the Fischer–Tropsch from coal was mostly composed of branched and cyclic alkanes that are difficult to speciate. For this fuel, we assumed that the distribution of branched and cyclic alkanes was similar to the distribution of *n*-alkanes in Fischer–Tropsch from natural gas (Corporan et al., 2011).

To overcome the limitations in the fuel composition data, we developed a mathematical relationship for C^* as a function of carbon number of a hydrocarbon to map the fuel composition data into the VBS. The relationship was derived using vapor pressure data from NIST for *n*-alkanes, cyclic alkanes, simple branched alkanes, single-ring aromatics, naphthalene, linear and cyclic alkenes, isoprene, and common terpenes (NIST, 2012). We found aromatics ($C^* = e^{(22.3 - carbon\#)/0.806}$) to have a slightly different relationship than alkanes and alkenes ($C^* = e^{(24.5 - carbon\#)/0.899}$).

The SOA formation for each fuel was characterized in a smog chamber. The experimental procedures are described in detail in Jathar et al. (2013). First, ammonium sulfate seed was added to the chamber to facilitate condensation and prevent nucleation of SOA products. Second, nitrous acid (HONO) was bubbled into the chamber. Third, the precursor mix (fuel) was introduced into the chamber using a heated septum. Finally, photo-oxidation was initiated by turning on the chamber UV black lights, which photolyzed HONO to produce OH radicals. NO and NO_2 formed as by-products of HONO irradiation resulted in a low VOC/NO_x ratio that was consistent with ratios found in urban polluted regions. The experiment was performed at low relative humidity (<5 %) and a temperature of around 298 K. The experiments were run for 2 to 7 h with an average OH exposure of 1.7×10^7 molecules h cm^{-3}; this corresponds to 17 h of photo-oxidation at an OH concentration of 10^6 molecules cm^{-3}. Therefore the data represent relatively fresh SOA formed under urban-like conditions.

Concentrations of gas-phase organic compounds were tracked using a GC-MS (Logue et al., 2009) and a proton-transfer reaction mass spectrometer (PTR-MS, Ionicon Analytik). The measured decay of individual species was used to estimate OH concentrations and OH exposure during the experiment. Particle-phase measurements were made using a scanning mobility particle sizer (SMPS, TSI Inc.) and a quadrupole or high-resolution aerosol mass spectrometer (Q-AMS or HR-AMS, Aerodyne Research Inc.) to measure non-refractory aerosol mass, size and composition. Both SMPS and Q-AMS/HR-AMS results were wall-loss corrected to calculate a lower and an upper bound on the total SOA formation (Hildebrandt et al., 2009). Depending on the experiment, an SOA measurement was characterized every 2 to 5 min.

Jathar et al. (2013) found that, for a unit mass of fuel reacted and at an atmospherically relevant OA concentration of $10\,\mu g\,m^{-3}$, unburned diesel formed the most SOA, followed by Jet Propellent-8 and Fischer–Tropsch from natural gas, gasoline and Fischer–Tropsch from coal. These trends reflected differences in fuel volatility and structure; fuels with higher carbon numbers and/or more aromatics formed more SOA. Fuels with lower carbon numbers and/or branched alkanes formed less SOA. In contrast to gasoline (Odum et al., 1997), the SOA formation from unburned diesel was not a strong function of the aromatic content. In fact, all diesels had similar SOA mass yields even though they varied in volatility and had different proportions of linear, branched and cyclic alkanes, alkenes and aromatics.

2.2 SOA models

We compared the measured SOA formation from Jathar et al. (2013) to SOA predictions/fits from a number of different models. In each model, fuel species react with the OH radical to form a set of semi-volatile surrogate products in the VBS (Donahue et al., 2006). The decay of the SOA precursor and production of the semi-volatile species is described using the following equations:

$$\frac{d[X_j]}{dt} = -k_{OH,X_j}[OH][X_j] \tag{1}$$

$$\frac{d[M_i|_{g+p}]}{dt} = \sum_j \alpha_{i,j} k_{OH,X_j}[OH][X_j], \tag{2}$$

where X_j is an individual SOA precursor j, $k_{OH,X}$ is the reaction rate constant between the oxidant [OH] and SOA precursor $[X_j]$, and $\alpha_{i,j}$ is the mass yield of product $M_i|_{g+p}$ ($g + p$ indicates gas plus particle phase) in VBS bin i. The index j indicates different precursors, either an individual fuel species or volatility bins of the fuel. The index i indicates different C^* bins in the VBS. The k_{OH} for each fuel species is listed in Tables S1 through S5 (in the Supplement). Gas-particle partitioning is calculated using Eq. (3).

$$\zeta_i = \left(1 + \frac{C_i^*}{C_{OA}}\right)^{-1} \tag{3}$$

$$C_{OA} = \sum_{i=1}^{N} \zeta_i \times M_i\big|_{g+p}\big|,$$

where ζ_i is the fraction of mass of $M_i\big|_{g+p}\big|$ in the particulate phase, C_i^* is the effective saturation concentration of M_i and C_{OA} is the total particulate OA concentration. As described in detail below, the different models use different lumping schemes for the fuel species and different mass yields $\alpha_{i,j}$.

We do not explicitly consider additional multi-generational oxidation of semi-volatile products because the SOA mass yields used in this work were derived from smog chamber experiments that spanned similar lengths in reaction time and OH exposure. Therefore, the SOA
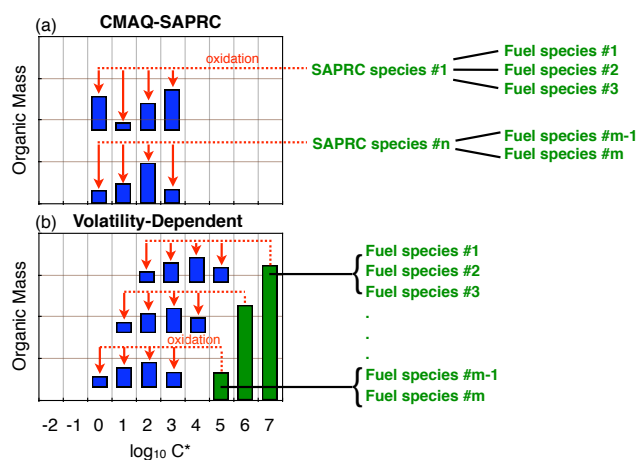

Figure 2. Schematics for the **(a)** CMAQ-SAPRC and **(b)** volatility-dependent SOA models.

mass yields already included some multigenerational aging, comparable to what occurred in our experiments.

We used two different types of models in this work, which are described in more detail in the subsequent sections. The first type was different versions of the SOA module used in the CMAQ model; they use volatility- and molecular structure-resolved schemes commonly found in SOA modules in CTMs. The second type was a volatility-dependent model that relates SOA production only to the precursor's volatility and ignores molecular structure.

2.2.1 CMAQ-SAPRC

Most SOA models used in CTMs account for SOA formation from speciated VOCs (isoprene, terpenes, single-ring aromatics and smaller alkanes and alkenes) using a set of lumped precursors that react to form products ($\alpha_{i,j}$) derived from fitting smog chamber data. The precursor lumping depends on gas-phase mechanisms, e.g., SAPRC or Carbon Bond. In this work, we used the SOA module implemented in CMAQ that employs the SAPRC gas-phase mechanism (Carlton et al., 2010). We refer to this as the CMAQ-SAPRC model, which is shown schematically in Fig. 2a.

The CMAQ-SAPRC model accounts for some differences in both precursor volatility and molecular structure. For example, it differentiates between alkanes (ALK), aromatics (ARO), isoprene (ISOP) and terpenes (TERP). For each molecular class/structure, there are typically multiple SAPRC lumped species to account for differences in reactivity with the OH radical. For a given molecular structure, reactivity partially correlates with volatility, so there is some differentiation with precursor volatility. There are few to no alkenes, terpenes or isoprene in any of the fuels; therefore to calculate SOA formation from the unburned fuel experiments involved accounting for four CMAQ-SAPRC lumped precursor species: ALK5, BENZ, ARO1, and ARO2. ALK5

is comprised of all alkanes larger than C_7. BENZ is benzene. ARO1 and ARO2 are lumped species for other aromatics.

SAPRC and most other gas-phase mechanisms were originally designed to simulate hydrocarbon–NO_x–ozone photochemistry, not SOA formation. Therefore, these lumping schemes do not account for some important differences in molecular structure and volatility that strongly influence SOA formation. For example, CMAQ-SAPRC does not differentiate between branched and linear alkanes even though branched alkanes have much lower SOA yields than equivalent-carbon linear alkanes (Lim and Ziemann, 2009; Tkacik et al., 2012). In addition, CMAQ-SAPRC also lumps all alkanes greater than C_7 into ALK5 even though alkanes smaller than about C_{10} form little SOA while alkanes larger than C_{10} have increasingly higher SOA yields (Lim and Ziemann, 2009; Presto et al., 2010; Tkacik et al., 2012). Therefore the SAPRC alkane lumping strategy has some shortcomings for simulating SOA formation.

Another reason that SAPRC and other mechanisms have not explicitly incorporated larger alkanes is that they are difficult to speciate using traditional GC-MS techniques and therefore rarely found in emission profiles and emission inventories (Robinson et al., 2007; Goldstein and Galbally, 2007). For example, in EPA's SPECIATE database (version 4.3) – the database of source profiles used to build most emission inventories for anthropogenic sources in the US – the emission profiles of the top six combustion sources of VOCs (on- and off-road gasoline, on- and off-road diesel, wood burning and open burning; they accounted for 97 % of all VOC combustion emissions in the US in 2008) have a negligible fraction (<1 %) of organics with a carbon number greater than 12. In contrast, numerous experimental studies have demonstrated that combustion sources emit substantial mass of organics with a carbon number greater than 12 (Schauer et al., 2002, 1999, 2001; Gentner et al., 2012; Presto et al., 2011). We have comprehensive speciation data for the fuel samples (Fig. 1); therefore we can explicitly incorporate larger organics into our analysis.

To investigate these potential shortcomings, the CMAQ-SAPRC model was run in three configurations that systematically increased the sophistication of the treatment of SOA precursors:

1. CMAQ-SAPRC (speciated): In this configuration, the model only considered organic compounds that have less than or equal to 12 carbons. This configuration was representative of the SOA treatment in most CTMs, including the standard treatment of SOA formation in CMAQ for alkanes and aromatics if one excludes multigenerational chemistry (Carlton et al., 2010). For Jet Propellent-8 and the Fischer–Tropsch experiments, we assumed that the branched and cyclic alkanes had the same carbon-number distribution as linear alkanes to determine the fraction of species that had fewer than or equal to 12 carbons. The lumping of the fuel

species (species with less than or equal to 12 carbons) into CMAQ-SAPRC model species used the standard SAPRC scheme, which was based on the reactivity of the species (Carter, 2007). This lumping is listed in Tables S1 through S5 in the Supplement. The SOA mass yields for the CMAQ-SAPRC model species were from the high NO_x pathway in Carlton et al. (2010), which were derived from fitting published smog chamber data. For the benefit of the reader, those SOA mass yields are listed in Table S6 in the Supplement. Those yields were refit to a four-bin VBS of $C^* = 1$, 10, 100 and $1000 \, \mu g \, m^{-3}$ to enable use with the VBS.

2. CMAQ-SAPRC (all): This version of the CMAQ-SAPRC model represented an extension of the base, CMAQ-SAPRC (speciated), model. In this version the model was extended to include all organics (speciated and larger/unspeciated), not just organics with 12 or fewer carbons. Again, the lumping of the fuel species was based on the reactivity of the species following the standard SAPRC scheme (Carter, 2007). This lumping is listed in Tables S1 through S5 in the Supplement. The larger alkanes (alkanes with more than 12 carbons) were added to the ALK5 lumped species (the largest alkane in the standard CMAQ-SAPRC mechanism) and the larger aromatics were added to the ARO2 lumped species. The CMAQ-SAPRC (all) model used the same SOA mass yields as the CMAQ-SAPRC (speciated) model; those yields are listed in Table S6 in the Supplement.

3. CMAQ-SAPRC (alkane-resolved): In this configuration, we extended the CMAQ-SAPRC (all) model to account for differences in SOA yields of alkanes with size and molecular-structure. This was done using the method of Pye and Pouliot (2012), which weights the emissions of different alkanes with their potential to form SOA. In other words, it accounted for the varying SOA potential of alkanes as a function of carbon number and linear/branched/cyclic structure to calculate an n-dodecane equivalent emission, which served as a surrogate for SOA formation from all alkanes. For reference, the weighting scheme of Pye and Pouliot (2012) is reproduced in Table S7. The SOA mass yields of n-dodecane used here were from Pye and Pouliot (2012) and are listed in Table S6 in the Supplement. The treatment of all non-alkane organics in CMAQ-SAPRC (alkane-resolved) was the same as in CMAQ-SAPRC (all).

To apply any of these versions of the CMAQ-SAPRC model to the fuel data required four SOA precursors (ALK5 or n-dodecane equivalent, BENZ, ARO1, ARO2), each of which required a four-parameter, two-product parameterization or a four-bin VBS parameterization (a total of 16 free parameters). Application of any of the CMAQ-SAPRC models did

not involve any fitting of the smog chamber data presented here.

2.2.2 Volatility-dependent model

In the volatility-dependent model, we assumed that SOA formation depended only on the precursor volatility. Therefore, this approach more explicitly accounted for the influence of volatility (or carbon number) on SOA formation than the CMAQ-SAPRC models, but it did not account for differences in molecular structure (e.g., alkane versus aromatic). The potential advantage of the volatility-dependent model was that it efficiently accounted for at least one property (volatility) known to influence SOA formation strongly.

The volatility-dependent model is shown schematically in Fig. 2b. Each precursor C^* bin (corresponding to X_j in Eq. 1) reacts with OH to form a distribution of semi-volatile products with mass yields $\alpha_{i,j}$. Each higher (or lower) precursor C^* bin is assumed to form the same product distribution but shifted by one C^* bin. This strategy is based on the work of Lim and Ziemann (2009b) and Presto et al. (2010), who found that for n-alkanes, the addition of two carbon atoms to an n-alkane shifted its corresponding SOA product distribution, on average, by one C^* bin.

Key inputs for the volatility-dependent model are the fuel volatility distribution and OH reaction rates for each precursor bin. Figure 1 shows the volatility distribution for each fuel. They were constructed from the composition data. The reaction rate of each precursor bin with the OH radical ($k_{OH,X}$) was derived from data in Atkinson and Arey (2003). We developed a mathematical relationship between C^* of a hydrocarbon and k_{OH}. Alkanes and aromatics have different relationships; alkanes: $k_{OH} = -1.84 \times 10^{-12} \log(C^*) + 4.27 \times 10^{-11}$ and aromatics: $k_{OH} = -5.7 \times 10^{-12} \log(C^*) + 1.14 \times 10^{-10}$.

To implement the volatility-dependent model in a CTM, one needs eight lumped precursor species, one for each C^* bin from $10^2 \, \mu g \, m^{-3}$ to $10^9 \, \mu g \, m^{-3}$. Since each precursor has the same five-bin VBS parameterization that is simply shifted in volatility space, the model only has five tunable parameters. As described below these parameters were derived by fitting the time-dependent experimental data.

3 Modeling results

3.1 CMAQ-SAPRC models

SOA predictions from the CMAQ-SAPRC (speciated) model were compared against measurements in Fig. 3a (Fig. 3 compares end-of experiment values; Fig. S1 in the Supplement shows additional experiment-by-experiment comparisons as a function of time.). This version of the CMAQ-SAPRC model displayed little skill in predicting the measured SOA formation ($R^2 = 0.04$, fractional error = 108 %). In addition to not predicting the variability in SOA forma-

tion, the CMAQ-SAPRC model over-predicted the Fischer–Tropsch from coal SOA by a factor of 5 and under-predicted the diesel SOA by a factor of 5. The CMAQ-SAPRC (all) model addressed one of the shortcomings of the CMAQ-SAPRC (speciated) model by accounting for all of the organics in the fuel. However, it only explained a little more of the variability ($R^2 = 0.21$, fractional error = 81 %) than the CMAQ-SAPRC (speciated) model (Fig. 3b; Fig. S2 in the Supplement shows additional experiment-by-experiment comparisons as a function of time). Like the CMAQ-SAPRC (speciated) model, it over-predicted the SOA formation from Fischer–Tropsch from coal. The minor improvement in the CMAQ-SAPRC (all) model resulted from slightly increased SOA predictions for the Jet Propellent-8 and the diesel fuels.

The under- and over-predictions were caused by shortcomings in the strategy used to lump organic precursors to model species. The CMAQ-SAPRC (all) and CMAQ-SAPRC (speciated) models used a very simple representation to model SOA formation from large alkanes, i.e., all alkanes with a carbon number greater than 7 are lumped into ALK5, which was assumed to form as much SOA as an n-undecane (C_{11}). This caused both of these models to over-predict the Fischer–Tropsch from coal SOA because most of its alkanes (88 % by mass centered around C_{11}) are branched. Branched alkanes have much lower SOA mass yields than linear and cyclic alkanes for the same carbon number (Lim and Ziemann, 2009; Tkacik et al., 2012). Likewise, both models under-predicted the diesel SOA because, on average, 80 % of its alkanes are larger than an n-undecane and 58 % of its alkanes are either linear or cyclic alkanes. Therefore, it is important that models account for both the carbon number and molecular structure if they are to predict SOA formation from alkanes accurately.

The CMAQ-SAPRC (alkane-resolved) model provided a more robust treatment of SOA formation from alkanes than the other two CMAQ-SAPRC models (Fig. 3c; Fig. S3 in the Supplement shows additional experiment-by-experiment comparisons as a function of time). Not surprisingly, it performed substantially better ($R^2 = 0.58$ and fractional error = 79 %) than the other two CMAQ-SAPRC models. Treating the alkanes with the scheme of Pye and Pouliot (2012) reduced the SOA formation from the branched alkanes in the Fischer–Tropsch fuels and increased the SOA formation from the linear and cyclic alkanes in the Jet Propellent-8 and diesel fuels. This underscores the importance of accounting for both carbon number and molecular structure of alkanes in SOA models.

3.2 Volatility-dependent model

In contrast to the CMAQ-SAPRC models, for the volatility-dependent model the time-dependent SOA data from all of the fuels were fit simultaneously to determine an SOA mass-yield matrix ($\alpha_{i,j}$) for the precursor mass distributed across a set of C^* bins; $\alpha_{i,j}$ is listed in Table 1. The
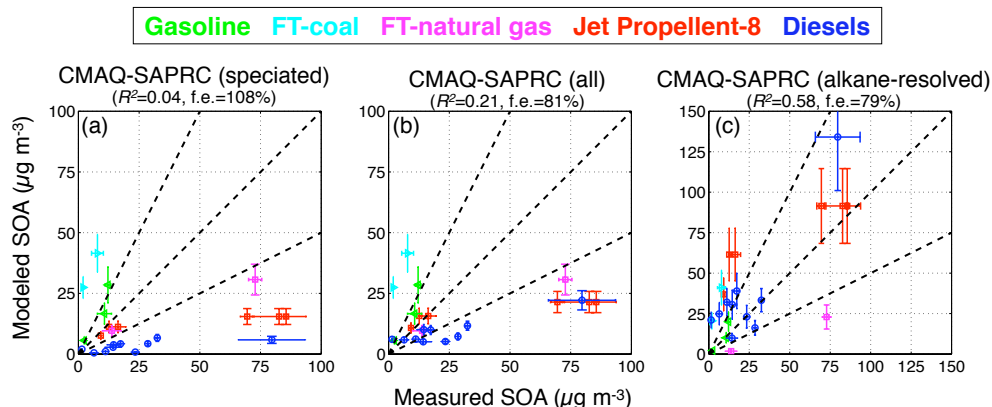
Figure 3. SOA predictions from the (**a**) CMAQ-SAPRC (speciated), (**b**) CMAQ-SAPRC (all) and (**c**) CMAQ-SAPRC (alkane-resolved) models compared to measurements. Panels compare end-of-experiment values; refer to Figs. S1–S4 in the Supplement that show experiment-by-experiment comparison as a function of time. The R^2 and fractional error values are listed in the parentheses. Fractional error $= \frac{1}{N}\sum_{i=1}^{N}\frac{|P-M|}{\frac{P+M}{2}}$; $P =$ Predicted, $M =$ Measured.

Table 1. SOA mass yields for model precursors in the volatility-dependent model.

Precursor C^*	Product C^* (μg m^{-3})												
	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	10^{0}	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}	10^{6}	10^{7}
10^{2} μg m^{-3}	0.011	0.078	0.034	0.006	0.297	–	–	–	–	–	–	–	–
10^{3} μg m^{-3}	–	0.011	0.078	0.034	0.006	0.297	–	–	–	–	–	–	–
10^{4} μg m^{-3}	–	–	0.011	0.078	0.034	0.006	0.297	–	–	–	–	–	–
10^{5} μg m^{-3}	–	–	–	0.011	0.078	0.034	0.006	0.297	–	–	–	–	–
10^{6} μg m^{-3}	–	–	–	–	0.011	0.078	0.034	0.006	0.297	–	–	–	–
10^{7} μg m^{-3}	–	–	–	–	–	0.011	0.078	0.034	0.006	0.297	–	–	–
10^{8} μg m^{-3}	–	–	–	–	–	–	0.011	0.078	0.034	0.006	0.297	–	–
10^{9} μg m^{-3}	–	–	–	–	–	–	–	0.011	0.078	0.034	0.006	0.297	–
10^{10} μg m^{-3}	–	–	–	–	–	–	–	–	0.011	0.078	0.034	0.006	0.297

goodness-of-fit is shown in Fig. 4a ($R^2 = 0.53$, fractional error $= 73\%$). Figure 4a also indicates that most predictions were within a factor of 2; Fig. S4 in the Supplement shows additional experiment-by-experiment comparisons as a function of time. This was comparable to the level of performance of the best CMAQ-based model, CMAQ-SAPRC (alkane-resolved), and much better than the other two CMAQ-based models. Therefore a model based on precursor volatility alone explained more than half of the variability in the measured SOA data. Of course, this model was fit to the data while the CMAQ-SAPRC models were based on published SOA parameterizations. However, the ability of the volatility-dependent model to describe the data highlights the strong dependence of precursor volatility on SOA formation.

There were several instances where the volatility-dependent model performance was poor. This likely was due to the influence of molecular structure. For example, although there was some variation in the measured diesel SOA mass, the volatility-dependent model predicted roughly the same SOA mass for the different diesel fuels because all of the diesel fuels had similar volatility distributions. The volatility-dependent model also predicted the same SOA

formation from Fischer–Tropsch from coal and Fischer–Tropsch from natural gas because they had the same volatility profile, but the measured yields were quite different. Fischer–Tropsch from coal is mostly composed of lower yield branched alkanes, while Fischer–Tropsch from natural gas is an equal mix of linear and branched alkanes.

In Fig. 4b we plot the SOA mass yields as a function of C^* from the volatility-dependent model and compare them to published single precursor yields. The comparisons are made at a typical atmospheric OA concentration of 5 μg m^{-3}. The SOA mass yields derived by fitting the volatility-dependent model are comparable to the single-compound data. For example, the fit for the volatility-dependent model predicted the range and drop in SOA mass yield for n-alkanes. Similarly, the fit predicted the right SOA mass yield range for alkenes and aromatics. The fit agreed with SOA mass yields for monoterpenes but under-predicts the SOA mass yields for isoprene and sesquiterpenes. It does this despite having no a priori information about the SOA mass yields for species in a given volatility range. This good agreement highlights the significant role that precursor volatility (carbon number/molecular size) has on SOA yields.
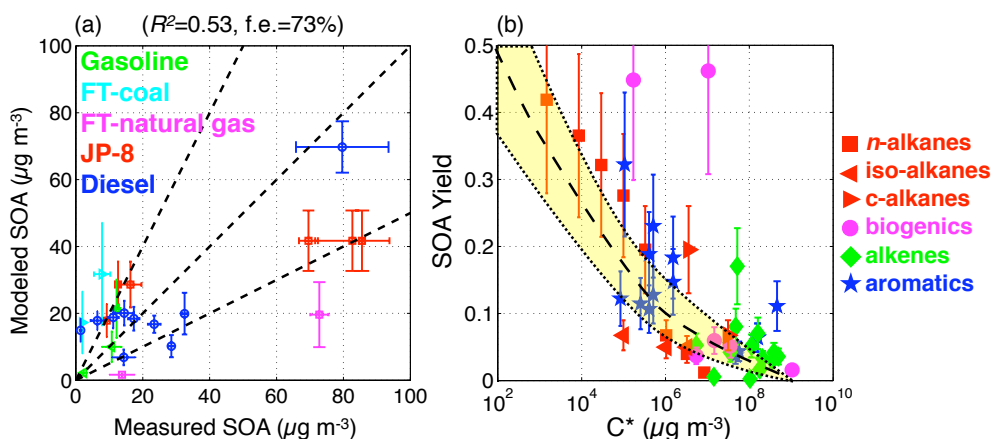
Figure 4. (a) SOA predictions from the volatility-dependent model compared to measurements. **(b)** SOA yield presented as a function of precursor C^* at a C_{OA} of $5\,\mu g\,m^{-3}$. The yellow band represents the lower and upper bound fits from the volatility-dependent model. n-alkane data are from Presto et al. (2010), cyclicalkane and iso-alkane data are from Tkacik et al. (2012), biogenic data are from Farina et al. (2010), alkene data are from Forstner et al. (1997), Na et al. (2006) and Keywood et al. (2004) and aromatic data are from Ng et al. (2007), Song et al. (2007), Hildebrandt et al. (2009), Chan et al. (2009) and Shakya et al. (2010). C^* values are determined either from the NIST database (NIST, 2012) or the EPA's Estimation Program Interface suite.

One could extend the volatility-dependent model to also incorporate molecular structure by having a different set of mass yields for different molecular structures. This would address some of the discrepancies shown in Fig. 4a. However, the fuel composition data were insufficient to determine a molecular structure-resolved volatility-dependent scheme.

4 Discussion

The comparisons of predictions from different versions of the CMAQ-SAPRC model with the SOA data underscored the importance of including all the speciated and larger/unspeciated organic precursor mass and accounting for the effects of alkane carbon number (molecular size) and molecular structure (linear/branched/cyclic) in SOA models. Historically, traditional SOA models employed in CTMs have only considered SOA formation from VOCs with a carbon number less than or equal to 12, mostly because larger/unspeciated SOA precursors are absent (or not properly accounted for) from emission profiles and inventories. These traditional SOA models have been sufficient to explain the SOA formation from single-ring aromatics present in unburned gasoline (Odum et al., 1997) but have failed in predicting the SOA formation from mixtures that contain larger/unspeciated organics. Addressing this need likely requires additional development of instruments and techniques to characterize these larger/unspeciated organics present in emissions. Furthermore, current ozone-focused VOC lumping schemes are inadequate for modeling SOA formation even if all of the organic mass is included in the model. The methods of Pye and Pouliot (2012) to determine an n-

dodecane equivalent mass were found to be both sufficient and efficient for modeling SOA formation from alkanes.

An SOA model based purely on the precursor volatility was able to capture the variability in SOA formation from different fuels that, similar to combustion emissions, contain complex mixtures of alkanes, alkenes and aromatics and their sub-types (linear, branched, cyclic, single-ring, multi-ring). In addition, the SOA yields of the volatility-dependent species (by C^*) were similar to data from single-compound experiments. The volatility-dependent approach implicitly assumed that the different fuels have a similar distribution of molecular structure. Hence, the SOA model based on volatility alone performed poorly when tested with SOA data from synthetic fuels like Fischer–Tropsch from coal and natural gas that have a much simpler composition, dominated by one class of species. For example, the Fischer–Tropsch from coal is mostly composed of branched alkanes. For these types of mixtures, it becomes important to account for the effects of molecular structure on SOA formation. For the same reason, the volatility-dependent approach may not be appropriate for modeling SOA formation from biogenic emissions (isoprene, monoterpenes, sesquiterpenes) because they have distinct molecular structures. This limitation of the volatility-dependent approach makes it unfit to predict changes in SOA formation with fuel reformulation that does not significantly change the volatility distribution of the fuel.

At present, the CMAQ-SAPRC-based models require 16 parameters (4 parameters each for 4 SOA precursors derived from previous smog chamber experiments) from high NO_x formation of SOA from alkanes and aromatics. In contrast, the volatility-dependent model required only 5 parameters (fit to unburned fuel smog chamber experiments) and had roughly the same skill as the CMAQ-SAPRC

model in predicting SOA formation from unburned fuel mixtures. Furthermore, the volatility-dependent model provides a compatible framework to include larger organic precursors on the basis of their volatility as and when they are characterized through emissions testing. These precursors would be tricky to incorporate into CMAQ-SAPRC-type models where lumping strategies are largely based on reactivity. So, although molecular structure influences SOA formation, given the results of this work, SOA formation as a function of volatility may be sufficient to model SOA formation in CTMs, especially from emissions of fossil-fuel sources. Biomass burning emissions also contain the same speciated SOA precursors as fossil fuel-based sources (Yokelson et al., 2013), so it is likely that a volatility-dependent approach would work to model SOA formation from biomass burning sources.

The Supplement related to this article is available online at doi:10.5194/acp-14-5771-2014-supplement.

Edited by: M. C. Facchini

References

Alnajjar, M., Cannela, B., Dettman, H., Fairbridge, C., Franz, J., Gallant, T., Gieleciak, R., Hager, D., Lay, C., Lewis, S., Ratcliff, M., Sluder, S., Storey, J., Yin, H., and Zigler, B.: Chemical and physical properties of the fuels for advanced combustion engines (FACE) research diesel fuels, Coordinating Research Council, 2010.

Carlton, A. G., Bhave, P. V., Napelenok, S. L., Edney, E. O., Sarwar, G., Pinder, R. W., Pouliot, G. A., and Houyoux, M.: Model representation of secondary organic aerosol in CMAQv4. 7, Environ. Sci. Technol., 44, 8553–8560, 2010.

Carter, W. P. L.: Development of the SAPRC-07 chemical mechanism and updated ozone reactivity scales, California Air Resources Board, Research Division, 2007.

Chan, A. W. H., Kautzman, K. E., Chhabra, P. S., Surratt, J. D., Chan, M. N., Crounse, J. D., Kürten, A., Wennberg, P. O., Flagan, R. C., and Seinfeld, J. H.: Secondary organic aerosol formation from photooxidation of naphthalene and alkylnaphthalenes: implications for oxidation of intermediate volatility organic compounds (IVOCs), Atmos. Chem. Phys., 9, 3049–3060, doi:10.5194/acp-9-3049-2009, 2009.

Corporan, E., Edwards, T., Shafer, L., DeWitt, M. J., Klingshirn, C., Zabarnick, S., West, Z., Striebich, R., Graham, J., and Klein, J.: Chemical, Thermal Stability, Seal Swell, and Emissions Studies of Alternative Jet Fuels, Energy Fuels, 25, 955–966, doi:10.1021/ef101520v, 2011.

De Gouw, J. A., Middlebrook, A. M., Warneke, C., Goldan, P. D., Kuster, W. C., Roberts, J. M., Fehsenfeld, F. C., Worsnop, D. R., Canagaratna, M. R., Pszenny, A. A. P., Keene, W. C., Marchewka, M., Bertman, S. B., and Bates, T. S.: Budget of organic carbon in a polluted atmosphere: Results from the New England Air Quality Study in 2002, J. Geophys. Res., 110, D16305, doi:10.1029/2004JD005623, 2005.

Donahue, N., Robinson, A., Stanier, C., and Pandis, S.: Coupled partitioning, dilution, and chemical aging of semivolatile organics, Environ. Sci. Technol., 40, 2635–2643, doi:10.1021/es052297c, 2006.

Dzepina, K., Cappa, C. D., Volkamer, R. M., Madronich, S., De-Carlo, P. F., Zaveri, R. A., and Jimenez, J. L.: Modeling the Multiday Evolution and Aging of Secondary Organic Aerosol During MILAGRO 2006, Environ. Sci. Technol., 45, 3496–3503, doi:10.1021/es103186f, 2010.

Forstner, H. J. L., Flagan, R. C., and Seinfeld, J. H.: Molecular speciation of secondary organic aerosol from photooxidation of the higher alkenes: 1-octene and 1-decene, Atmos. Environ., 31, 1953–1964, 1997.

Fraser, M. P., Cass, G. R., Simoneit, B. R. T., and Rasmussen, R.: Air quality model evaluation data for organics. 4. C_2–C_{36} non-aromatic hydrocarbons, Environ. Sci. Technol., 31, 2356–2367, 1997.

Gentner, D. R., Isaacman, G., Worton, D. R., Chan, A. W., Dallmann, T. R., Davis, L., Liu, S., Day, D. A., Russell, L. M., Wilson, K. R., Weber, R., Guha, A., Harley, R. A., and Goldstein, A. H.: Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions, Proc. Natl. Acad. Sci., 109, 18318–18323, 2012.

Goldstein, A. H. and Galbally, I. E.: Known and unexplored organic constituents in the earth's atmosphere, Environ. Sci. Technol., 41, 1514–1521, doi:10.1021/es072476p, 2007.

Grieshop, A. P., Logue, J. M., Donahue, N. M., and Robinson, A. L.: Laboratory investigation of photochemical oxidation of organic aerosol from wood fires 1: measurement and simulation of organic aerosol evolution, Atmos. Chem. Phys., 9, 1263–1277, doi:10.5194/acp-9-1263-2009, 2009.

Hildebrandt, L., Donahue, N., and Pandis, S.: High formation of secondary organic aerosol from the photo-oxidation of toluene, Atmos. Chem. Phys., 9, 2973–2986, doi:10.5194/acp-9-2973-2009, 2009.

Jathar, S., Farina, S., Robinson, A., and Adams, P.: The influence of semi-volatile and reactive primary emissions on the abundance and properties of global organic aerosol, Atmos. Chem. Phys., 11, 7727–7746, doi:10.5194/acp-11-7727-2011. 2011.

Jathar, S., Miracolo, M., Presto, A., Donahue, N., Adams, P., and Robinson, A.: Modeling the formation and properties of traditional and non-traditional secondary organic aerosol: problem formulation and application to aircraft exhaust, Atmos. Chem. Phys., 12, 9025–9040, doi:10.5194/acp-12-9025-2012, 2012.

Jathar, S. H., Miracolo, M. A., Tkacik, D. S., Donahue, N. M., Adams, P. J., and Robinson, A. L.: Secondary Organic Aerosol Formation from Photo-Oxidation of Unburned Fuel: Experimental Results and Implications for Aerosol Formation from Combustion Emissions, Environ. Sci. Technol., 47, 12886–12893, doi:10.1021/es403445q, 2013.

Lim, Y. B., and Ziemann, P. J.: Effects of molecular structure on aerosol yields from OH radical-initiated reactions of linear,

branched, and cyclic alkanes in the presence of NO_x, Environ. Sci. Technol., 43, 2328–2334, 2009.

Logue, J., Huff-Hartz, K., Lambe, A., Donahue, N., and Robinson, A.: High time-resolved measurements of organic air toxics in different source regimes, Atmos. Environ., 43, 6205–6217, 2009.

May, A. A., Levin, E. J. T., Hennigan, C. J., Riipinen, I., Lee, T., Collett, J. L., Jimenez, J. L., Kreidenweis, S. M., and Robinson, A. L.: Gas-particle partitioning of primary organic aerosol emissions: 3. Biomass burning, J. Geophys. Res. Atmos., 118, 11327–11338, doi:10.1002/jgrd.50828, 2013a.

May, A. A., Presto, A. A., Hennigan, C. J., Nguyen, N. T., Gordon, T. D., and Robinson, A. L.: Gas-particle partitioning of primary organic aerosol emissions: (2) Diesel vehicles, Environ. Sci. Technol., 47, 8288–8296, 2013b.

May, A. A., Presto, A. A., Hennigan, C. J., Nguyen, N. T., Gordon, T. D., and Robinson, A. L.: Gas-particle partitioning of primary organic aerosol emissions: (1) gasoline vehicle exhaust, Atmos. Environ., 77, 128–139, 2013c.

Ng, N., Kroll, J., Chan, A., Chhabra, P., Flagan, R., and Seinfeld, J.: Secondary organic aerosol formation from m-xylene, toluene, and benzene, Atmos. Chem. Phys, 7, 3909–3922, doi:10.5194/acp-7-3909-2007, 2007.

NIST Chemistry WebBook: http://webbook.nist.gov/chemistry/, 2012.

Odum, J. R., Jungkamp, T., Griffin, R., Flagan, R. C., and Seinfeld, J. H.: The atmospheric aerosol-forming potential of whole gasoline vapor, Science, 276, 96–99, 1997.

Pankow, J. F.: An absorption model of gas/particle partitioning of organic compounds in the atmosphere, Atmos. Environ., 28, 185–188, 1994.

Presto, A. A., Miracolo, M. A., Donahue, N. M., and Robinson, A. L.: Secondary organic aerosol formation from high-NO x photooxidation of low volatility precursors: n-alkanes, Environ. Sci. Technol., 44, 2029–2034, 2010.

Presto, A. A., Nguyen, N. T., Ranjan, M., Reeder, A. J., Lipsky, E. M., Hennigan, C. J., Miracolo, M. A., Riemer, D. D., and Robinson, A. L.: Fine particle and organic vapor emissions from staged tests of an in-use aircraft engine, Atmos. Environ., 45, 3603–3612, 2011.

Pye, H., and Seinfeld, J.: A global perspective on aerosol from low-volatility organic compounds, Atmos. Chem. Phys, 10, 4377–4401, doi:10.5194/acp-10-4377-2010, 2010.

Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. A., Sage, A. M., Grieshop, A. P., Lane, T. E., Pierce, J. R., and Pandis, S. N.: Rethinking organic aerosols: Semivolatile emissions and photochemical aging, Science, 315, 1259–1262, 2007.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T.: Sources of fine organic aerosol. 2. Noncatalyst and catalyst-equipped automobiles and heavy-duty diesel trucks, Environ. Sci. Technol., 27, 636–651, 1993.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T.: Sources of fine organic aerosol. 9. Pine, oak, and synthetic log combustion in residential fireplaces, Environ. Sci. Technol., 32, 13–22, 1998.

Saunders, S. M., Jenkin, M. E., Derwent, R. G., and Pilling, M. J.: Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile organic compounds, Atmos. Chem. Phys., 3, 161–180, doi:10.5194/acp-3-161-2003, 2003.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T.: Measurement of emissions from air pollution sources. 2. C_1 through C_{30} organic compounds from medium duty diesel trucks, Environ. Sci. Technol., 33, 1578–1587, 1999.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T.: Measurement of Emissions from Air Pollution Sources. 3. C_1–C_{29} Organic Compounds from Fireplace Combustion of Wood, Environ. Sci. Technol., 35, 1716–1728, 2001.

Schauer, J. J., Kleeman, M. J., Cass, G. R., and Simoneit, B. R. T.: Measurement of Emissions from Air Pollution Sources. 5. C_1–C_{32} Organic Compounds from Gasoline-Powered Motor Vehicles, Environ. Sci. Technol., 36, 1169–1180, 2002.

Shakya, K. M., and Griffin, R. J.: Secondary Organic Aerosol from Photooxidation of Polycyclic Aromatic Hydrocarbons, Environ. Sci. Technol., 44, 8134–8139, doi:10.1021/es1019417, 2010.

Shrivastava, M. K., Lane, T. E., Donahue, N. M., Pandis, S. N., and Robinson, A. L.: Effects of gas particle partitioning and aging of primary emissions on urban and regional organic aerosol concentrations, J. Geophys. Res.-Atmos., 113, D18301, doi:10.1029/2007JD009735, 2008.

Song, C., Na, K., Warren, B., Malloy, Q., and Cocker III, D. R.: Secondary organic aerosol formation from the photooxidation of p-and o-xylene, Environ. Sci. Technol., 41, 7403–7408, 2007.

Tkacik, D. S., Presto, A. A., Donahue, N. M., and Robinson, A. L.: Secondary organic aerosol formation from intermediate-volatility organic compounds: cyclic, linear, and branched alkanes, Environ. Sci. Technol., 46, 8773–8781, doi:10.1021/es301112c, 2012.

Tsimpidi, A., Karydis, V., Zavala, M., Lei, W., Molina, L., Ulbrich, I., Jimenez, J., and Pandis, S.: Evaluation of the volatility basis-set approach for the simulation of organic aerosol formation in the Mexico City metropolitan area, Atmos. Chem. Phys, 10, 525–546, doi:10.5194/acp-10-525-2010, 2009.

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

Yokelson, R. J., Burling, I. R., Gilman, J., Warneke, C., Stockwell, C. E., Gouw, J. D., Akagi, S., Urbanski, S., Veres, P., and Roberts, J. M.: Coupling field and laboratory measurements to estimate the emission factors of identified and unidentified trace gases for prescribed fires, Atmos. Chem. Phys., 13, 89–116, doi:10.5194/acp-13-89-2013, 2013.