We thank the reviewers for the insightful comments and have made every effort to address them appropriately. A point-by-point response is listed below. Reviewer comments are in black, our response is in italic and blue and changes made to the manuscript are in maroon.

Based on the reviewer comments, we have made three changes to the models that we would like to bring to your attention before you read the response to individual comments. This is mostly because the names and some of their implementation in the ACPD manuscript and the final manuscript are different.
1. We have removed the Empirical model. The original motivation for the Empirical model came from the Jathar et al. (2013) paper where the fuel data were fit to facilitate comparison with the emissions data. None of the current-day SOA models in CTMs are based on the Empirical model (i.e., SOA is modeled directly as a reaction product of non-methane organic emissions) so we did not see any reason to keep it.
2. We have changed the name and some of the implementation of the Traditional models. In the revised manuscript, we refer to them as CMAQ-SAPRC models and bring them more in line with the implementation in most chemical transport models and climate models. The CMAQ-SAPRC models are built based on the SOA implementation in CMAQ (Community Multiscale Air Quality model) that utilizes the gas-phase mechanism SAPRC (Carlton et al., 2010;Pye and Pouliot, 2012). The CMAQ-SAPRC (speciated) model only includes organic species that have less than 12 carbons. The CMAQ-SAPRC (all) model includes all organic species. The CMAQ-SAPRC (alkane-resolved) model accounts for the varying SOA potential of alkanes with carbon number and structure (linear/branched/cyclic). Sections 2.2.1 and 3.1 have changed substantially.
3. The name of the Volatility-based model is changed to Volatility-Dependent model.

Response to reviewer comments.
Reviewer 1

1. In the title and in the introduction, the authors state quite clearly that they are investigating the effect of molecular structure on SOA formation. However, I was quite disappointed when I read that n-alkanes, isoalkanes and cycloalkanes end up being lumped into ALK4, ALK5 based on the SAPRC scheme. The volatility based scheme also does not account for molecular structure. In other words, the shift in volatility as a result of oxidation should depend on fragmentation mechanisms and molecular structure, but that effect is ignored when they are all lumped together. There is a differentiation between aromatics and alkanes, but not molecular structure of the backbone (degree and position alkyl branching, rings etc.). Therefore I do not believe that the manuscript in its current form is investigating molecular structures directly. I could be misunderstanding the authors, so I would like to see a clarification if I am mistaken. Currently, I believe the message is “molecular structure does not matter for SOA from complex fuels, but it does for fuels with simpler composition”. If so, I believe that is an important message to bring out, and the title should be changed to reflect that.

Based on this and the other reviewer’s comments, we have substantially changed the implementation of the ‘Traditional’ models. Instead of using the SOA model in PMCAMx, we now use the SOA model in CMAQ that runs with the SAPRC gas-phase mechanism. In the revised manuscript, we refer to the models as CMAQ-SAPRC models. In the CMAQ-SAPRC *(speciated) and CMAQ-SAPRC (all) implementations, all SOA forming alkanes are lumped into the model species ALK5; this emulates their treatment in most chemical transport models. However, in the CMAQ-SAPRC (alkane-resolved) model, we use the work of Pye and Pouliot (2012) to account for the varying SOA yields of alkanes as a function of carbon number and structure. We weigh the emissions of alkanes with their potential to form SOA and calculated an n-dodecane equivalent emission, which is then used to model SOA formation from all alkanes. The SOA dependence of aromatics on volatility and molecular structure is kept the same across the CMAQ-SAPRC models, i.e. benzene is treated separately, ARO1 are C7 and C8 aromatics that have higher SOA yields and ARO2 are C8+ aromatics that have lower SOA yields. With this change, our work evaluates the treatment of volatility and molecular structure in the current generation of CMAQ-SAPRC models on SOA formation. For a detailed description of the CMAQ-SAPRC models, refer to Section 2.2.1. Also, as suggested by the reviewer, we have changed the manuscript title to better reflect the focus on the paper*: "Testing secondary organic aerosol models using smog chamber data for complex precursor mixtures: Influence of precursor volatility and molecular structure".

To some degree, the CMAQ-SAPRC models account for both volatility and molecular structure but not very explicitly. The Volatility-based (now the Volatility-Dependent model) explicitly accounts for the influence of volatility but ignores molecular structure. Despite that, its model performance is similar to the best CMAQ-SAPRC model. The reviewer has not misunderstood our claim and we have highlighted it better in the revised manuscript in the 'Discussion' section.

"An SOA model based purely on the precursor volatility was able to capture the variability in SOA formation from different fuels, that like 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*) are similar to data from single compound experiments. The volatility-dependent approach implicitly assumes that the different fuels have a similar distribution of molecular structure. Hence, the SOA model based on volatility alone performs 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 to model SOA formation from biogenic emissions (isoprene, monoterpenes, sesquiterpenes) because they have distinct molecular structures."

2. One caveat of this study is that the results apply only to urban SOA formation, where reactivities are dominated by alkanes, simple alkenes, and aromatic compounds, or compounds of fossil fuel origins. Readers should be cautioned that the trends in SOA yields are unlikely to be the same for biogenic compounds, like isoprene, and monoterpenes, or biomass burning emissions. Therefore it is quite ambitious to say that volatility-based frameworks are sufficient for all CTMs, as stated in the last paragraph.

We agree with the reviewer that we have only modeled SOA formation from SOA precursors that are representative of fossil fuel emissions and therefore the results only apply to fossil fuel *sources. We have clarified this point in the revised manuscript. However, biomass burning emissions also contain the same speciated SOA precursors as fossil fuel-based emissions (Yokelson et al., 2013), so it is likely that the results from this work would apply to biomass burning emissions too. In Figure 3(b), the volatility-based parameterization seems to reproduce the SOA yields (at a* C_{OA} *of 5 μg* m^{-3}*) for monoterpenes but under-predict the SOA yields for isoprene and sesquiterpenes. Since biogenic SOA precursors are much more unique (*C_5*,* C_{10} *or* C_{15}*) than their anthropogenic counterparts (spread over* C_6*-*C_{25}*), it would be more appropriate to model biogenic SOA precursors explicitly than to use a volatility-based model. We have added the following text to the 'Discussion' section:*

"The volatility-dependent approach implicitly assumes that the different fuels have a similar distribution of molecular structure. Hence, the SOA model based on volatility alone performs 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 to model SOA formation from biogenic emissions (isoprene, monoterpenes, sesquiterpenes) because they have distinct molecular structures."

"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. 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."

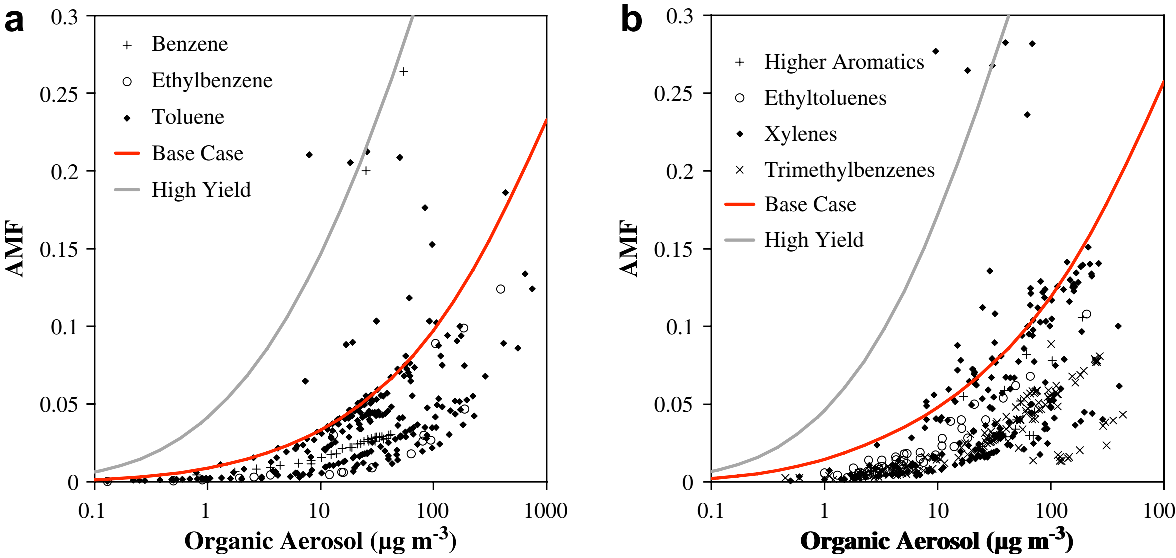
3. The average OH rate constant is 1E-11, which seems low. The species that will be primarily responsible for SOA formation will likely be larger alkanes or aromatics, which will have rate constants of at least 2 or 3E-11 (naphthalene/heptadecane). So it would make sense to me that multiple generations are indeed more relevant. (For 3E-11 rate constant, and OH exposure of 1.7E7 molec cm-3 hr, that would be almost 2 lifetimes of oxidation.) Also, Ziemann et al. and the Caltech group showed that multiple generations are important for SOA formation. Therefore, one should take multiple generations into account. It is entirely reasonable to expect that the fractional bias and fractional error would be greater in such case (because the model will be more complex), but this is more realistic and accurate.

The reviewer is correct that a higher OH rate constant of the precursor and its subsequent products would result in substantial multigenerational aging that could potentially influence SOA formation. However, there are two reasons why we do not consider multigenerational aging in our models. First, the SOA mass yields used in the CMAQ-SAPRC models have been derived using smog chamber experiments that span similar lengths in reaction times and OH exposure. Therefore, the SOA mass yields already include some multi-generational aging, comparable to what occurred in our experiments. Second, the experiments by the Ziemann and the Caltech group were done at much higher OH exposures (at least a factor of 2 higher) than our experiments. Therefore the amount of multigenerational aging in our experiments is less than *that in the Ziemann and CalTech experiments. We have modified the text in the manuscript to reflect the reviewer's concern.*

"We do not explicitly consider additional multi-generational oxidation of semi-volatile products for two reasons. First, the SOA mass yields used in this work were derived that smog chamber experiments that span similar lengths in reaction time and OH exposure. Therefore, the SOA mass yields already include some multigenerational aging, comparable to what occurred in our experiments."

4. In table S5, diesel seem to contain a disproportionately large amount of C9 aromatic (one order of magnitude higher than other species). That would suggest to me that SOA would be dominated by those from these C9 aromatics (trimethylbenzenes?). Is that common knowledge among scientists studying SOA formation? (There seem to a lot more focus on benzene / toluene / xylenes, judging from Ng et al., 2006).

The reviewer is correct to point out that diesel has a large fraction of C9 aromatics and that the bulk of the SOA is coming from C9 aromatics. However, these diesels were designed to span a wide range in fuel properties and may not be typical of commercial diesel. Diesel #9, which is closest to the commercially-available diesel, has 15% of its mass as C9. Despite the dominance of C9 aromatics, it is not concerning from an SOA perspective. In a figure compiled by Lane et al. (2008) (extracted below), C7, C8, C9 and C9+ aromatics seem to have similar SOA yields (or aerosol mass fraction (AMF)). So, although recent SOA studies have focused on C6, C7 and C8 aromatics, the SOA yields from these might be applicable to higher aromatics.



Reviewer 2

1. Motivation. On 24409L2–4, the authors state "Given the success of volatility-based schemes, it could be worthwhile to explore their capabilities for modeling all of SOA, not just SOA from S/IVOCs." This statement runs counter to the body of literature that the authors just summarized, in which single-compound studies demonstrated that SOA formation from low-volatility compounds has a strong dependence on molecular structure (24407L13–26). Moreover, it runs counter to the authors' own earlier work (Jathar et al., 2013) which showed that SOA yields from volatile fuels (e.g., gasoline and Fischer Tropsch from natural gas) are more sensitive to chemical composition than to fuel volatility. The authors proceed to construct a "Volatility-based model" which uses mass yields that are independent of molecular structure (Table 2). As expected, this model fails to match the observed variation in SOA quantities produced from fuels with similar volatility but differing composition (24420L23 – 24421L5). In their revised manuscript, I would urge the authors to use the volatility-based model merely to illustrate the shortcomings of ignoring molecular structure and focus more effort on improving the accuracy of the "Traditional" model.

Based on the reviewer's comment, we have used variants of the CMAQ-SAPRC model and refocused the paper much more on describing the performance of the Traditional or CMAQ-SAPRC models. These models vary in their capacity to account for lower-volatility organics and the influence of volatility and molecular structure of the precursors on SOA formation.

The CMAQ-SAPRC models seem to account for both volatility and molecular structure but not very explicitly, e.g. all aromatics greater than xylene, irrespective of volatility, are lumped into ARO2. In contrast, the Volatility-Dependent model accounts for at least one property (volatility) known to strongly influence SOA formation. The model performance for the Volatility-Dependent model (when fit) is similar to the best CMAQ-SAPRC (alkane-resolved) model and explains about half the variability in the observed SOA formation. Further, 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 that are resolved on the basis of reactivity. So, although molecular structure influences SOA formation, given the results of this work, SOA formation as a function of volatility may be a sufficient substitute to CMAQ-SAPRC models to simulate SOA formation in CTMs.

We agree with the reviewer that molecular structure is an important attribute to consider while modeling SOA formation and we have revised the text in the 'Discussion' section to talk about the shortcomings of the Volatility-Dependent model.

"The volatility-dependent approach implicitly assumes that the different fuels have a similar distribution of molecular structure. Hence, the SOA model based on volatility alone performs 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 to model SOA formation from biogenic emissions (isoprene, monoterpenes, sesquiterpenes) because they have distinct molecular structures."

2. On 24417L4, the authors state that the yield parameters for the Empirical model were fit using the entire dataset and that Fig. 2 shows only the level of agreement at the end of each experiment. From this, I interpret that each photo-oxidation experiment lent multiple data points to be fit with the models. At what time resolution were the experimental data supplied for model fitting?

Yes, each photo-oxidation experiment contributed multiple data points to be fit with the models. Depending on the experiment, an SOA mass concentration measurement was interpreted every 2 to 5 minutes. Since most of the SOA growth leveled off in about 20 to 40 minutes, each experiment provided a maximum of 4 to 20 meaningful data points to be fit with the model. We have added the following sentence to Section 2.1 (Experimental data): "Depending on the experiment, an SOA measurement was interpreted every 2 to 5 minutes."

On a related note, we have also added an experiment-by-experiment comparison of model predictions with measurements for the four different models in the supporting material (Figure S1 through S4).

3. The authors suggest that the high fractional error is due to experimental variability and measurement uncertainty (24417L15-20). This explanation leaves much to be desired. For the values which appear to be poorly fit at the end of an experiment (e.g., low-concentration JP-8 experiments and several diesel experiments), was the model performance better at earlier stages of the photo-oxidation? When computing the best-fit parameters, were all time points of the experiment given equal weight? Was standard least-squares fitting applied or was some effort made to place more importance on fitting the atmospherically-relevant (i.e., low) concentrations?

Based on the reviewer's comments, we thought hard about the purpose of the Empirical model in this paper (and elsewhere in the research community) and decided to remove it from the manuscript. The original motivation for the Empirical model came from the Jathar et al. (2013) paper where the fuel data were fit to facilitate comparison with the emissions data. None of the current-day SOA models in CTMs are based on the Empirical model (i.e., SOA is modeled directly as a reaction product of non-methane organic emissions) so we did not see any reason to keep it. Instead, we have now focused our manuscript on the performance of different versions of the CMAQ-SAPRC models and a volatility-dependent model, all of which are currently used in some form in CTMs.

We have investigated the temporal performance of the model (see figures below). We did not find the model performance to be better at earlier stages of photo-oxidation, i.e. the model was biased high or low for the entire duration of the experiment. Hence, a standard least-square method was employed by giving each data point the same weight. We also investigated fitting all of the diesel experiments together to arrive at a single fit for all the diesels. This fit reproduces nine of the ten diesel experiments within a factor of two, which is within the uncertainty typical *of smog chamber experiments, which investigate SOA formation from photo-oxidation. To cite a few examples, Hildebrandt et al. (2009) found that the SOA mass yields for toluene varied by a factor of two. When compared against Ng et al. (2007), the SOA mass yields of Hildebrandt et al. (2009) were a factor of two higher. Similarly, SOA mass yields for similar experiments on naphthalene from Chan et al. (2009) and Shakya and Griffin (2010) vary by at least a factor of two.*

The different diesels differ modestly in their composition, so we fit each diesel fuel separately. By doing so, we find that except for two JP-8 experiments and one diesel experiment, the Empirical model reproduced the measured SOA very well (see figure below; R^2*=0.86, fractional error=30%). The cause of these outliers is not known; potential explanations include large sensitivity of the SOA formation to gas/particle wall-loss rates, oxidant or radical concentrations or VOC-to-NOx ratios.*

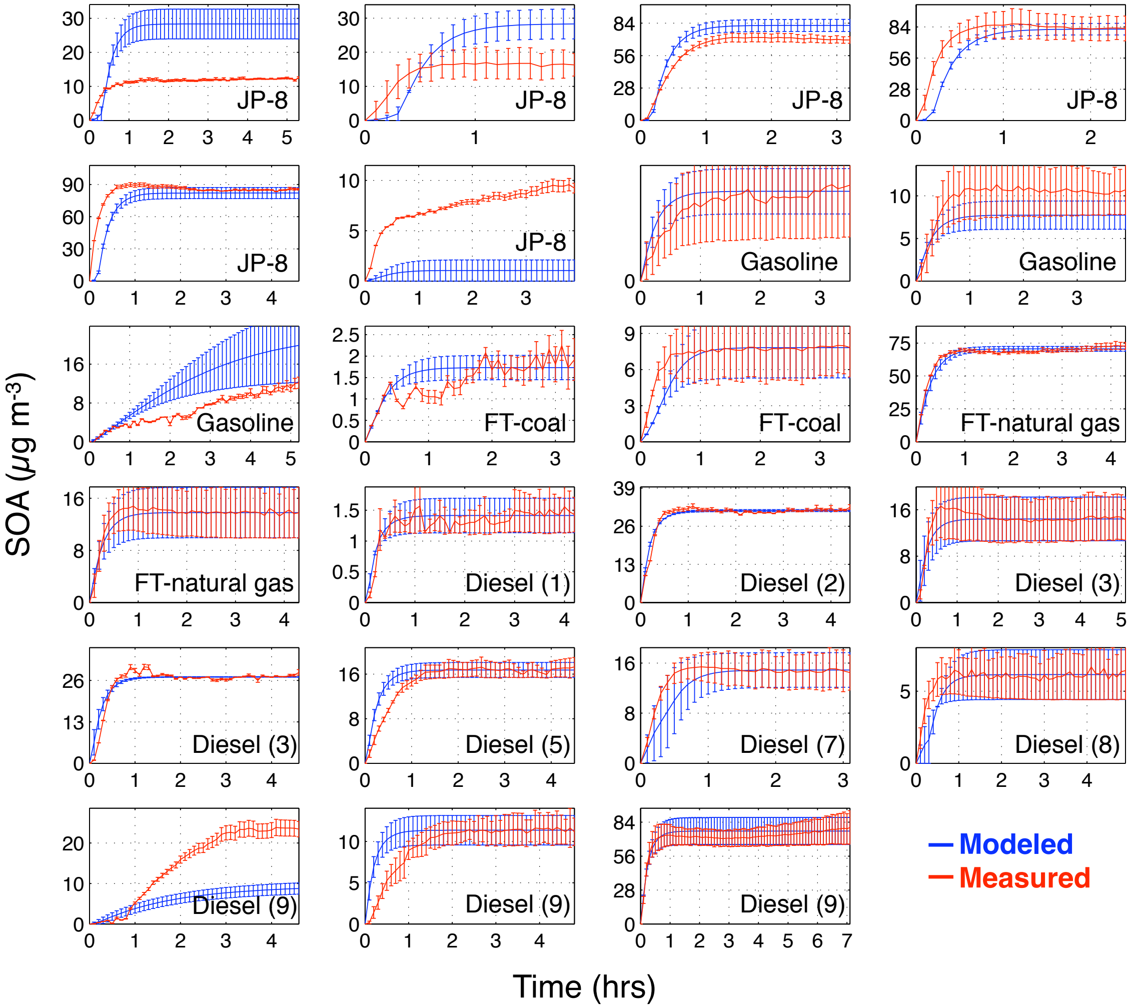


Figure: Empirical model predictions of SOA compared against measurements as a function of time for 23 experiments from Jathar et al. (2013).

5. Are the yield parameters for the Empirical model in this paper the same as given in Table 1 of Jathar et al. (2013) and depicted by the solid lines in Fig. 3 of that paper?

Yes, the yield parameters for the Empirical model are the same as in Table 1 of Jathar et al. (2013). Since the Empirical model was removed in the revised manuscript, Table 1 was removed too.

6. In the revised manuscript, I would prefer to see more analysis of the model performance throughout a given experiment and more exploration of plausible hypotheses for why the best possible model can result in a complete mismatch with the experimental data. For example, perhaps the Empirical model can be parameterized such that each precursor class (e.g, aromatic, branched alkane, straight alkane) yields a unique combination of volatility-binned products rather than all fuel species yielding the same products. If the temporal resolution of the experimental data is 2 or 4 minutes as indicated by Jathar et al. (2013), there should be enough data to fit more than 5 parameters per fuel.

As we mentioned earlier, the revised manuscript does not include the Empirical model. For the CMAQ-SAPRC models and the Volatility-Dependent model, the revised manuscript does include an experiment-by-experiment comparison between model-predictions and measurements in the supporting material (Figures S1 through S4). In those figures, one can see that for all the models and most of the experiments, the model predictions were consistently biased either low or high for the entire duration of the experiment. Further, in most experiments the majority of the SOA was formed within the first 20 to 40 minutes of the experiment. The consistent bias in model predictions and the rapid SOA production means that the temporal information does not provide much additional information for the modeling effort.

The reviewer's concern about the Empirical model's fit being poor is addressed in point number (3) above.

7. For many readers, the term "Traditional" (Section 2.3.2) will likely be misinterpreted as the 2-product approach used historically for modeling SOA yields. However, the authors use the term to imply traditional groupings of precursors (e.g., ALK5, OLE2, ARO1) but then assign yields to those lumped species using a newer volatility-based scheme (e.g., Lane et al., 2008). The potential for misinterpretation increases further when one reads in Section 2.3.3 that the third model is referred to as "Volatility-based." To minimize confusion, I suggest referring to the second model type as something like "SAPRC lumping" rather than "Traditional" and the third model type perhaps as "Structure-independent" rather than "Volatility-based."

To be clear, the 2-product scheme and the volatility basis set approach, when fit to the same SOA chamber data and in the absence of any aging schemes, are identical. In the revised manuscript, the "traditional" model is replaced by the CMAQ-SAPRC model that uses SAPRC to lump precursors into model species (ALK5, BENZ, ARO1, ARO2) and uses the two-product parameterization from Carlton et al. (2010) to model SOA formation.

We agree with the reviewer that the nomenclature of the different models, in this work and those found elsewhere, is bound to cause confusion. To communicate our results better, we have *decided to use the SOA model in CMAQ (Community Multiscale Air Quality model) with SAPRC lumping as our base model instead of our 'Traditional' model. The SOA model in CMAQ with SAPRC lumping (hereon referred to as the CMAQ-SAPRC model) is reflective of the SOA treatment in most chemical transport models and climate models.*

We find that the reviewer's suggestion of using the name 'structure-independent' model potentially confusing because it could also mean that it is 'everything-else-dependent' instead of just being dependent on volatility. To avoid confusion between the 'volatility basis set' and 'volatility-based model', we have changed the name of our model from 'volatility-based model' to 'volatility-dependent' model.

8. Later in this section, the authors refer to a "speciated" and "all" version of the Traditional SOA model. On 24415L26-27, they also make reference to a "Traditional (base)" model. In Section 3.2, the authors introduce a "Traditional (extended)" model. On 24415L24 and L27, the authors state that the SAPRC lumping is listed in Tables S1 through S5. However, those tables do not mention "Traditional (all)." It seems that the authors modified their methodology midway through the study but then failed to revise the manuscript in a consistent manner. In the revised manuscript, the authors ought to describe all of the models in Section 2.3 as well as some brief description of the genetic optimization algorithm that is brought up later in the manuscript.

We apologize for the oversight. The 'Traditional (base)' model in the supporting material should have been the 'Traditional (all)' model. As mentioned earlier, we have changed the names of models to CMAQ (speciated), CMAQ (all) and CMAQ (alkane-resolved). In the revised manuscript, we have ensured that they are described in detail in Section 2.2 and their names are consistent in the supporting material. Since we do not refit any of the CMAQ-SAPRC models, we did not use the genetic algorithm.

9. On 24416L3, the authors state that their Traditional model had 45 free parameters (9 SAPRC precursors × 5 volatility bins). This is an outright misrepresentation. First, this number includes parameters related to isoprene and terpenes which the authors stated on 24414L14-15 that they would no longer discuss in this manuscript. Second, the Table S6 where these parameters are tabulated include only 4 volatility bins per precursor at the high-NOx conditions relevant to this manuscript. Third, these aren't "free" parameters in the present study. As stated on 24414L10-11, they were determined prior to this study by fitting chamber data on single compounds. This misrepresentation pervades the Abstract and should be fixed in the revised manuscript. On 24416L23-24, the authors state that the VBS only had 5 free parameters. But Table 2 lists a much larger number of unique mass yields. Moreover, the values in Table 2 do not exactly reflect the schematic in Fig. 1c (i.e., values are not equal along each diagonal).

The reviewer is correct to point out that our 'free parameter' comparison between the CMAQ-SAPRC models and the Volatility-Dependent model was not justified, specially considering that we were only looking at high NO_x SOA formation from fossil-fuel based emissions. By 'free' or 'tunable' parameters, we intend to say that, those are the number of parameters that are needed to describe SOA formation from organic emissions now matter how they were fit. In other words, *irrespective of how you derive those parameters, a CMAQ-SAPRC model would need 16 parameters but a Volatility-Dependent model would need only 5. We have edited the text in the revised manuscript as follows:*

Section 2.2.1: "The SOA mass yields for the CMAQ-SAPRC model species are from the high NOx 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…. To apply any of these versions of the CMAQ-SAPRC model to the fuel data requires 4 SOA precursors (ALK5 or n-dodecane-equivalent, BENZ, ARO1, ARO2) each of which requires a 4 parameter two-product parameterization (a total of 16 tunable parameters). Application of any of the CMAQ-SAPRC models did not involve any fitting of the smog chamber data presented here."

Section 2.2.2: "To implement the Volatility-Dependent model in a CTM, one needs 8 lumped precursor species, one for each C* bin from 10^2 µg m^{-3} to 10^9 µg m^{-3}. Since each precursor has the same 5-bin VBS parameterization that is simply shifted in volatility space, the model only has 5 tunable parameters. As described below these parameters were derived by fitting the experimental data."

In Table 2, one finds a much larger number of unique parameters because the parameters were deliberately accommodated to occupy the VBS space between a C of 0.1 and a C* of 1000 µg m^{-3}. Since this might cause confusion, the table has been changed to:*

"Table 1: SOA mass yields for model precursors in the Volatility-Dependent model"

Precursor C* (µg m^{-3})	Product C* (µg m^{-3})												
	1E-5	1E-4	1E-0	1E-2	1E-1	1E+0	1E+1	1E+2	1E+3	1E+4	1E+5	1E+6	1E+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

10. Apparent bias against Traditional (speciated) model. The authors state on 24418L11 that, for fuels other than gasoline, "only 30 – 50% of the mass was considered in the Traditional (speciated) model." But on 24411L15-17, the authors state that the hydrocarbons other than n-alkanes were assumed to have the same carbon number distribution as the n-alkanes in that fuel. Given this assumption, all of the isoalkanes in the Fischer Tropsch (natural gas) should have been mapped to ALK5 and produced some SOA in the Traditional (speciated) model. But Table S3 reveals that the isoalkanes in this fuel were not mapped to any of the SAPRC species. It appears that the authors made structural model assumptions that bias their result toward better performance for the "Traditional (all)" than the "Traditional (speciated)" model.

In the earlier version of the manuscript, we had only included the speciated organics in the Traditional (speciated) model and therefore not included the unspeciated 'isoalkanes' and 'cycloalkanes' present in the Fischer-Tropschs in that model. Since almost all of the isoalkanes and cycloalkanes in the Fischer-Tropschs are smaller than a carbon number of 12, we could ideally assume that those would be speciated using conventional GCMS techniques. Based on the reviewer's comment, we have included the isoalkane and cycloalkane mass in the Fischer-Tropschs in the CMAQ-SAPRC (speciated) model (earlier Traditional (speciated) model) based on their proportion of n-alkanes that are smaller than a carbon number of 12. The change is reflected in the text as shown below and in Table S2 and Table S3 in the supporting material.

"For Jet Propellent-8 and the Fischer-Tropschs, 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 less than 12 carbons."

11. Perhaps the most immediately useful finding of the present study is buried on 24419L5–7. As noted by the authors, the over-prediction of gasoline SOA by a factor of 5 is a clear indication that the mass yields for aromatics in PMCAMx are overestimated. This result should be highlighted in the Abstract. Another insightful result is reached on 24419L8-17. Here, the authors conclude that "it is important to differentiate between branched, straight, and cyclic alkanes when modeling SOA formation." Pye and Pouliot (ES&T 2012) reached the same conclusion and developed a parameterization to treat alkanes of varying length and structure within a SAPRC-lumped framework. Perhaps their SOA parameters should be selected in the Traditional model for this study rather than using the less-refined alkane parameters of Murphy and Pandis (2010) or applying the genetic optimization algorithm.

Based on the reviewer's suggestion, in the revised manuscript, we have refrained from using the SOA model in PMCAMx, built and tested a new model (CMAQ-SAPRC (alkane-resolved)) based on the findings of Pye and Pouliot (2012) and not explored further improvement using the genetic optimization algorithm.

In our modeling with the CMAQ-SAPRC models, we found that the gasoline data were not over-predicted because of the high SOA mass yields for aromatics. The gasoline data were over-predicted because the CMAQ-SAPRC (speciated and all) model formed too much SOA from the small alkanes. By employing the scheme in Pye and Pouliot (2012) in the CMAQ-SAPRC (alkane-resolved) model, the over-prediction in the gasoline and Fischer-Tropsch from coal data was resolved.

Minor Comments

12. Section 2 makes extensive reference to a document by Jathar et al. (2013) but this does not appear in the References list. Simon et al. (2010) is also missing from the References list, so I think the list should be checked for other omissions. 24410L15-17. Based on Figure 3 by Jathar et al. (2013), I believe this statement is conditional on the total concentration of organic PM

(COA). For example, the pink and red lines in that figure crossover at high COA and the SOA yields from certain fuels are undetermined at low COA. This paragraph ought to be reworded to mitigate the chances of being taken out of context in the future.

We apologize for the oversight. The Simon et al. (2010) reference was no longer required. Jathar et al. (2013) have been added to the references list. To address the reviewer's concern about being taken out of context, we have reworded the sentence to:

"Jathar et al. (2013) found that, for a unit mass of fuel reacted and an atmospherically-relevant OA concentration of 10 µg 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."

13. 24413L1. Tables S1 – S5 ought to be condensed into 2 tables. One table can provide the composition of all fuels (Mass %) and the other table can provide kOH and lumping information for all compounds. The current format contains a lot of duplicate information and is prone to error. For example, Tables S3 and S4 provide different kOH values for the same species (n-octane).

Consolidating Tables S1-S5 would be hard since the description of the fuel composition was different for each fuel. We have retained the same structure (a table for each fuel) and updated the lumping based on the new CMAQ-SAPRC models.

14. 24413L3 – formatting error. Should be a vertical bar after Mi to match with Eqs. 2 and 4.

The equations have been corrected.

15. 24413L4-11 – please indicate how many hours was a typical exposure. The only indication I can find is in Fig. 3 by Jathar et al. (2013), which shows data from 4 hours of photo-oxidation. While the information about e-folding lifetimes is insightful, it's an inadequate substitute for reporting actual reaction times.

We have changed the manuscript to add detail about reaction times.

Section 2.1: "The experiments were run for 2 to 7 hours with an average OH exposure of 1.7 x 10^{7} molecules-hr cm^{-3}; this corresponds to 17 hours 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."

16. 24413L7 – units of reaction rate are incorrect; should be cm3 instead of cm−3

The text has been corrected.

17. Figure 1 is a very helpful complement to the text descriptions given for the 3 SOA models. I offer a couple of recommendations to make the schematics even clearer. Species #1 is listed above species #n to the right of Fig. 1a, and the positions are reversed beside Fig. 1b. Unless there is a reason for this switch, it would be clearer to put both in the same order. The blue bars in Figs. 1a and 1b are aligned by their lower edges but in Fig. 1c they are aligned to the upper edges. I believe the depiction would be clearer if all bars were aligned along their bottom edges.

The figures have been changed to reflect the comment.

24414L4 – change “is” to “are”

The text has changed and we did not have to make this correction.

24414L7 – change “i.e.” to “e.g.” since these are examples of the preceding statement and not a rephrasing of the same.

The text has changed and we did not have to make this correction.

24415L5 – I suggest omitting the word “exclusively” because that can be interpreted to mean the SPECIATE database has only one use.

The word ‘exclusively’ was removed.

If possible, Equations (5) and (6) should be moved to the caption of Fig. 4 so they don’t receive undue emphasis.

We no longer use Equation (5) since we discuss the under- and over-prediction in the text. Equation (6) has been moved to the caption of Figure 3.

References

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, Environmental Science & Technology, 44, 8553-8560, 2010.

Chan, A. W. H., Kautzman, K. E., Chhabra, P. S., Surratt, J. D., Chan, M. N., Crounse, J. D., Kurten, A., Wennberg, P. O., Flagan, R., and Seinfeld, J. H.: Secondary organic aerosol formation from photooxidation of naphthalene and alkylnaphthalenes: implications for oxidation of intermediate volatility organic compounds (IVOCs), Atmospheric Chemistry and Physics, 9, 3049-3060, 2009.

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

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, Environmental Science & Technology, 47, 12886-12893, 10.1021/es403445q, 2013.

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.

Pye, H. O., and Pouliot, G. A.: Modeling the role of alkanes, polycyclic aromatic hydrocarbons, and their oligomers in secondary organic aerosol formation, Environmental Science & Technology, 46, 6041-6047, 2012.

Shakya, K. M., and Griffin, R. J.: Secondary Organic Aerosol from Photooxidation of Polycyclic Aromatic Hydrocarbons, Environmental Science & Technology, 44, 8134-8139, doi:10.1021/es1019417, 2010.

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, Atmospheric Chemistry and Physics, 13, 89-116, 2013.