Atmospheric
Chemistry
and Physics
Discussions

Open Access

Interactive comment on "Modeling the influence of precursor volatility and molecular structure on secondary organic aerosol formation using evaporated fuel experiments" *by* S. H. Jathar et al.

Anonymous Referee #2

Received and published: 24 November 2013

The present study builds upon recent experimental work by the same investigators (Jathar et al., 2013) in which they measured the SOA formation from eleven different unburned fuels across a broad range of organic PM concentrations ($<$ 1 to almost 100 μg/m^3). That study documented a number of important observations. For example, SOA yields from volatile fuels are more sensitive to aromatic content than carbon number whereas SOA yields from heavier fuels are less sensitive to aromatic content. The present study takes the next logical step by attempting to replicate these observations using a mathematical model. If successful, such a model could be used to simulate the atmospheric SOA formation from various fuels.

The manuscript is novel and addresses scientific questions that are significant and relevant to *ACP*. However, the scientific quality and presentation need more work. The motivation for exploring the capability of volatility-based schemes to represent all SOA is weak. The methodological description is quite difficult to follow and not sufficiently complete to allow reproduction by other scientists. Finally, some of the most important results gained from this analysis are not highlighted in the Abstract (rather, they are buried in Section 3.2) and I believe that even more useful insights can be drawn from this rich dataset than what the authors present. Although some portions of the manuscript are very well-written, a number of weaknesses prevent me from recommending publication at this time.

Major Comments

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.

Methodology is incomplete and prone to misinterpretation.
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?

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?

In the photo-oxidation of JP-8 on 11/8/11, for example, the authors measured 9.6 ± 0.6 μg/m^3 of SOA at the end of the experiment whereas the empirical model yielded no SOA at all. The authors need to make a greater effort to explain results such as this before they conclude that the Empirical model presents an "upper limit on the performance of other models."

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?

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

C9384

5 parameters per fuel.

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."

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.

On 24416L3, the authors state that their Traditional model had 45 free parameters (9 SAPRC precursors $\times$ 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.

C9385

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).

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.

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.

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

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 k_{OH} 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 k_{OH} values for the same species (n-octane).

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

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.

24413L7 – units of reaction rate are incorrect; should be cm^3 instead of cm^{-3}

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.

24414L4 – change "is" to "are"

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

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

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