**Responses to Referee #1**

In this work, Murphy et al. present the v5.2 of the Community Multiscale Air Quality (CMAQ) model, which takes into account the semivolatile phase partitioning and photochemical aging of POA emissions, and its evaluation over multiple seasons and locations. In addition, they have introduced into the model a new surrogate species that represents the potential SOA from combustion emissions (pcSOA) and explore its sensitivity to the emission scaling factor and the oxidation rate constant. Overall, the manuscript is very well written and the presentation is clear. However, I have a major concern regarding the implementation of pcSOA and the way that is presented. I recommend this study for publication after considering the following comments.

We thank the reviewer for their careful consideration of our manuscript and sharing constructive feedback. We have accepted their revisions in almost every case and are grateful for the added clarity they will provide. Where we disagree with the reviewer, we have explained our rationale fully in the responses below, and we look forward to continued dialogue.

General comments:

1. The authors implemented pcSOA into CMAQv5.2 by using the SIMPLE parameterization for SOA formation in polluted urban regions Hodzic and Jimenez (2011). The optimal empirical parameters used for the formation of pcSOA (emission factors and oxidation reaction rate constants) are based on the observed proportionality of SOA concentrations and excess CO downwind of urban locations (i.e., Mexico City and Pasadena) under low biomass burning conditions. Therefore, pcVOC emissions includes all the anthropogenically emitted organic vapors and pcSOA expresses all the anthropogenic SOA. However, CMAQ already includes emissions of traditional anthropogenic VOCs and emissions of SVOCs (e.g., VSVPO3) and IVOCs (i.e., VIVPO1) in the gas phase, and therefore, the formation of SOA from anthropogenic VOCs, IVOCs, and SVOCs is currently double counted.

The reviewer is correct in their description of the SIMPLE parameterization and its basis in proportionality of SOA to excess CO. However, the parameters we have chosen for this study do not correspond exactly to those used previously. First, the emissions of the VOC precursor in Hodzic and Jimenez (2011) were scaled to CO emissions; in this study, they are scaled to POA emissions. For our base case scaling factor, we began with the optimal parameter from Hayes et al. (2013), 0.069 g VOC (g CO)$^{-1}$, and we converted it to g VOC (g POA)$^{-1}$ using the average ratio of CO to POA concentrations predicted by the model in the California domain. For the California scenarios, this combination of parameters worked well. For the larger domain, as we show in the manuscript, the existing measurements suggested that we decrease the emission scale factor by 32%. The take-home point is that we are not applying the parameters from Mexico City and Pasadena studies a priori to our cases. Instead, we are using a similar method after accounting for all of the known sources that CMAQ has included in the past. In this way, pcSOA only accounts for potentially missing sources and pathways.

2. The first author of the manuscript have proposed a very effective nomenclature for atmospheric organic aerosols (Murphy et al., 2014). I would expect some consistency and I suggest to the authors to follow the naming convention of Murphy et al. (2014). In fact, the progress in the field of organic aerosols has come along with many complex and inconsistent abbreviates, therefore, adopting a standardized nomenclature will benefit the effective communication of the results to

the scientific community. However, this is only a stylistic preference and it is up to the authors to adopt it.

We appreciate the reviewer's suggestion and did of course use that previous publication as the starting point for our choice of names in the current scheme. As it stands, there are pre-existing constraints in CMAQ for naming chemical species (e.g. all particle-phase species begin with "A", aqueous SOA compounds already existed in the model, biomass burning was not considered separately here, etc). Additionally, we are not making use of an important component of the Murphy et al. (2014) scheme, the tracking of volatility at the point of emission. In future versions of the model, we anticipate that more elements of that scheme will become accessible to us and help standardize the nomenclature to a greater extent.

Specific comments:

1. I find the introduction section very well written. Please consider including the following recent articles as well: Page 2 lines 24-25: (Shiraiwa et al., 2017); Page 3 line 10: (Ots et al., 2016; Jathar et al., 2017); Page 3 lines 19-20: (Ma et al., 2016); Page 4 line 1: (Tsimpidi et al., 2017).

The references have been added to the manuscript as the reviewer suggests.

2. Page 5 line 25: The Robinson et al. (2007) distribution includes emission factors up to the $10^6$ µg $m^{-3}$ volatility bin. However, CMAQv5.2 only have species with saturation concentration up to $10^4$ µg $m^{-3}$. Can you please include into the manuscript (maybe in Table 1?) the emission factors used for each of the primary organic material?

The emission factors have been added to table 1.

**Table 1.** Properties of semivolatile POA species in CMAQv5.2.

| Species Name (Particle/Vapor) | Mol. Weight g $mol^{-1}$ | $C^{*,1}$ µg $m^{-3}$ | $N_C$ | O:C | $H_{vap}$ kJ $mol^{-1}$ | POA Emission Frac.[2] |
|---|---|---|---|---|---|---|
| ALVPO1/VLVPO1~~[2]~~[3] | 218 | 0.1 | 13.0 | 0.185 | 89 | 0.09 |
| ASVPO1/VSVPO1~~[3]~~[4] | 230 | 1 | 14.5 | 0.123 | 85 | 0.09 |
| ASVPO2/VSVPO2 | 241 | 10 | 16.0 | 0.073 | 81 | 0.14 |
| ASVPO3/VSVPO3 | 253 | 100 | 17.5 | 0.032 | 77 | 0.18 |
| AIVPO1/VIVPO1~~[4]~~[5] | 266 | 1000 | 19.0 | 0.00 | 73 | 0.5 |
| ALVOO1/VLVOO1~~[5]~~[6] | 136 | 0.01 | 5.0 | 0.886 | 93 | NA |
| ALVOO2/VLVOO2 | 136 | 0.1 | 5.5 | 0.711 | 89 | NA |
| ASVOO1/VSVOO1 | 135 | 1 | 6.0 | 0.567 | 85 | NA |
| ASVOO2/VSVOO2 | 135 | 10 | 6.5 | 0.447 | 81 | NA |
| ASVOO3/VSVOO3 | 134 | 100 | 7.0 | 0.345 | 77 | NA |

[1] $C^*$ values are defined at reference temperature 298 K.

[2] Robinson et al. (2007)

[3] A = Aerosol; V = Vapor; LV = Low Volatility; PO = Primary Organic

[4] SV = Semivolatile

[5] IV = Intermediate Volatility

[6] OO = Oxidized Organic

3. Page 6 lines 10-15: Does the oxidation of SVOC result in products with both higher and lower volatility than the precursor? I would suggest adding a reaction as an example and reporting the stoichiometric coefficients that you are using for each of the products.

Yes, each reaction takes into account fragmentation consistent with Eq. 1. Thus products of higher and lower volatility are produced. We have added an example reaction and its stoichiometry to the supplemental information. Further, we have added the entire stoichiometry for the aging mechanism to the supplemental information.

> “Table S1 shows the stoichiometric coefficients derived from the branching ratio formula (Eq. 1) and the aging kernel published in Donahue et al. (2012). As an example, we provide the oxidation stoichiometry of the oxidation of VSVPO2, one of the surrogates for primary semivolatile vapors:
>
> $$\begin{aligned}\text{VSVPO2} + \text{OH} \rightarrow\ & \text{OH} + 0.3856 * \text{VLVPO1} + 0.0950 * \text{VSVPO1} \\ & + 0.1373 * \text{VSVPO2} + 0.0005 * \text{VSVPO3} \\ & + 0.2051 * \text{VLVOO1} + 0.1764 * \text{VLVOO2}\end{aligned} \qquad \text{Eq. S2}$$
>
> These molar yields are derived in order to conserve carbon. Hydroxyl radicals are assumed to be regenerated by the oxidation reactions so that the aging of POA does not perturb the oxidant budget; the overall reactivity associated with POA mass is small compared to that associated with VOCs. The stoichiometric coefficients indicate that low O:C surrogates (VLVPO1-VIVPO1) in addition to high O:C surrogates (VLVOO1-VSVOO3). This feature ensures that the evolution of bulk O:C behaves similarly to the more detailed approaches. This is described in more detail in the following section.”

4. Section 2.2: Do VOCs participate to only one oxidation step or are subject to multigenerational aging? Either way, I believe that SOA from anthropogenic VOCs is double counted due to the presence of pcSOA (see above in the first general comment). Hodzic and Jimenez (2011) do not include any traditional anthropogenic VOC species (e.g., benzene) other than the lumped surrogate anthropogenic specie VOCA, which is analogous to the pcVOC used here.

The secondary compounds formed from VOC oxidation are subject to only the first oxidation step (i.e. no multigenerational aging) consistent with Griffin et al. (1999). We acknowledge the reviewer’s point and concede that it is possible that some of the anthropogenic SOA from multigenerational aging of VOCs is “double-counted” by the model configuration. However, the total first-generation contribution of SOA from anthropogenic VOCs predicted by the model at Pasadena, for example, is minor (9.0%) compared to the contribution of pcSOA predicted (67.6%). We have added a speciated plot depicting the average contributions of individual OA categories to the total in the supporting information (Fig. S4). This simple comparison illustrates the significant gap between the first-generation anthropogenic SOA yields in CMAQ and observations of urban SOA dominated by vehicles as well as other sources. Although we have taken great care in this study to recommend a set of robust parameters for pcSOA formation, this assessment relies on data from specific time periods, like 2011 over the continental U.S. We emphasize that these parameters will be revised in the future as better-constrained parameters for anthropogenic OA formation, like multigenerational SOA yields, volatility-resolved emission factors, and heterogeneous reaction rates, etc, are implemented. These improvements will enable a more consistent bottom-up understanding of the urban OA problem, but until then, this top-down approach provides useful predictions of bulk OA formation and exposure.

5. Page 7 line 3: I am not sure that the name “pcVOC” is the more appropriate to use since this species also includes unspeciated SVOC/IVOC emissions.

Although the reviewer’s point is well-taken, we prefer to keep the existing name for this surrogate species. We consider the term VOC as it used here to be a more appropriate term based on how it is implemented in the model, because partitioning of this species, even in very cold (e.g. upper troposphere) or very concentrated (e.g. biomass burning plumes) conditions is not considered. We agree with the reviewer about the possibility of confusion here, so we have added a statement to the methods section 2.3.

> “A new surrogate VOC species (potential VOC from combustion emissions, pcVOC) is introduced with an emission rate that is scaled to the POA mass emission rate. This species does not partition directly to the particle phase and in that respect, behaves as a VOC in the model. It:is oxidized with OH to form a low volatility condensable vapor, potential secondary organic gas from combustion emissions, pcSOG (table 2).”

6. Page 7 line 25: The saturation concentration reported here (10-3 _g m-3) is not the same with the value reported in Table 2 (10-5 _g m-3).

Updated.

7. Page 11 line 32: The sum of the fractions reported here is 99%.

The reported percentages in the original study sum to 99% rather than 100% due to rounding error. Please see Liu et al. (2012).

8. Page 12 line 4: Please add the fractional biases of your nonvolatile POA simulation for comparison.

Updated.

9. Page 19 line 31: Please consider deleting the “from anthropogenic and biogenic carbon sources”.

We have made the change.

10. Figures 1, 2 and 3: Please add in the figures captions what the boxes and the whiskers represent.

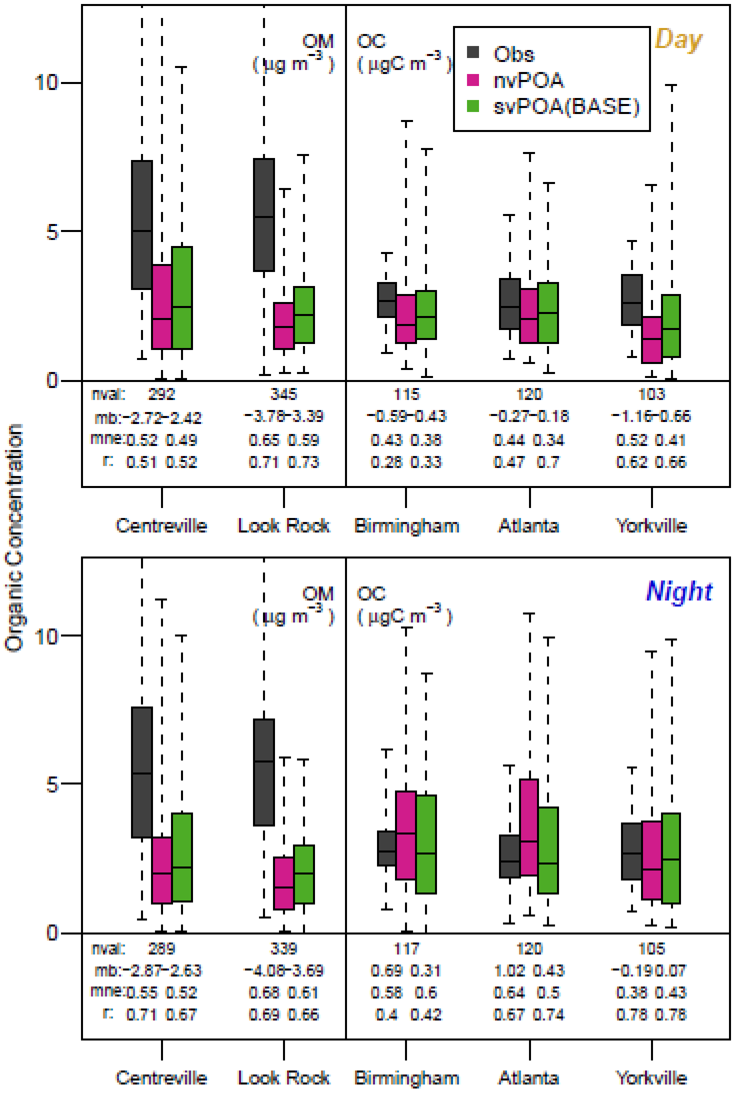
We have added this information to the Figure captions.

> “The boxes denote the 25th and 75th percentiles of each dataset while the whiskers extend to the most extreme points.”

11. Figure 2: According to the figure caption, the observed values over Birmingham, Atlanta and Yorkville are organic carbon (in μgC $m^{-3}$). Are the modelled values converted to OC as well? I would suggest converting everything (both observed and modelled values) to OA (in μg $m^{-3}$) in

order to be consistent with the values over Centreville and Look Rock, which are shown in the same figure, as well as with the title of the y axis.

The model values are in μgC m$^{-3}$ for Birmingham, Atlanta, and Yorkville. Because we view the application of OM:OC ratios from the model to be more appropriate than using uncertain OM:OC ratios for the field data, we prefer to keep the comparisons that are in OC units. We also point out that the comparison across sites is not the focus of this analysis. Instead, we seek the most reliable comparison of modeled to measured data at each site independently. We agree that the y-axis is confusing, and we have updated the figure to be more clear on this point.



12. Figure 3. Please add here (and discuss in the text) the model performance for POA and OOA over Centreville, Look Rock and Sacramento where there are also AMS measurements available.

Unfortunately, HOA/OOA factors for these three sites are not available. The PMF analyses for Sacramento does not exist to our knowledge. The analyses at Centreville and Look Rock were able to isolate a small number of factors including an aged OA, an isoprene-derived OA and a biomass burning OA factor. However, the relative contributions of these factors would not have been affected significantly by the improvements made in this study, which are related to the anthropogenic combustion OA. Displaying these factors in Fig. 3 would also not be appropriate since none of them correspond to the HOA factor already present. We have added a statement to the manuscript explaining why these comparisons are absent.

> "Figure 3 compares CMAQ OA species to AMS factors derived from PMF analysis at Pasadena, Bakersfield, and Cool. PMF analysis was not available at the Sacramento site, and the Centreville and Look Rock sites showed negligible presence of HOA throughout the SOAS campaign."

13. Figure 5: I would recommend deleting the boxes and the whiskers behind each trend and keeping only the lines. This figure is too overloaded and it is very difficult for the reader to take any extra-information other than the trends of the lines.

Updated

References

Hodzic, A. and Jimenez, J. L.: Modeling anthropogenically controlled secondary organic aerosols in a megacity: a simplified framework for global and climate models, Geosci. Model Dev., 4, 901-917, 2011.

Jathar, S. H., Woody, M., Pye, H. O. T., Baker, K. R., and Robinson, A. L.: Chemical transport model simulations of organic aerosol in southern California: model evaluation and gasoline and diesel source contributions, Atmos. Chem. Phys., 17, 4305-4318, 2017.

Ma, P. K., Zhao, Y., Robinson, A. L., Worton, D. R., Goldstein, A. H., Ortega, A. M., Jimenez, J. L., Zotter, P., Prévôt, A. S. H., Szidat, S., and Hayes, P. L.: Evaluating the impact of new observational constraints on P-S/IVOC emissions, multi-generation oxidation, and chamber wall losses on SOA modeling for Los Angeles, CA, Atmos. Chem. Phys. Discuss., 1-35, 2016.

Murphy, B. N., Donahue, N. M., Robinson, A. L., and Pandis, S. N.: A naming convention for atmospheric organic aerosol, Atmos. Chem. Phys., 14, 5825-5839, 2014.

Ots, R., Young, D. E., Vieno, M., Xu, L., Dunmore, R. E., Allan, J. D., Coe, H., Williams, L. R., Herndon, S. C., Ng, N. L., Hamilton, J. F., Bergström, R., Di Marco, C., Nemitz, E., Mackenzie, I. A., Kuenen, J. J. P., Green, D. C., Reis, S., and Heal, M. R.: Simulating secondary organic aerosol from missing diesel-related intermediate-volatility organic compound emissions during the Clean Air for London (ClearfLo) campaign, Atmos. Chem. Phys., 16, 6453-6473, 2016.

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.

Shiraiwa, M., Li, Y., Tsimpidi, A. P., Karydis, V. A., Berkemeier, T., Pandis, S. N., Lelieveld, J., Koop, T., and Poschl, U.: Global distribution of particle phase state in atmospheric secondary organic aerosols, Nature Communications, 8, 2017.

Tsimpidi, A. P., Karydis, V. A., Pandis, S. N., and Lelieveld, J.: Global-scale combustion sources of organic aerosols: Sensitivity to formation and removal mechanisms, Atmos. Chem. Phys. Discuss., 2017, 1-38, 2017.

## References for Response

Griffin, R. J., et al. (1999). "Estimate of global atmospheric organic aerosol from oxidation of biogenic hydrocarbons." Geophysical Research Letters **26**(17): 2721-2724.

Hodzic, A. and Jimenez, J. L.: Modeling anthropogenically controlled secondary organic aerosols in a megacity: a simplified framework for global and climate models, Geosci. Model Dev., 4, 901-917, 2011.

Murphy, B. N., Donahue, N. M., Robinson, A. L., and Pandis, S. N.: A naming convention for atmospheric organic aerosol, Atmos. Chem. Phys., 14, 5825-5839, 2014.