Responses to the Reviews of "Characterization of Organic Aerosol across the Global Remote Troposphere: A comparison of ATom measurements and global chemistry models"

Anonymous Referee #3

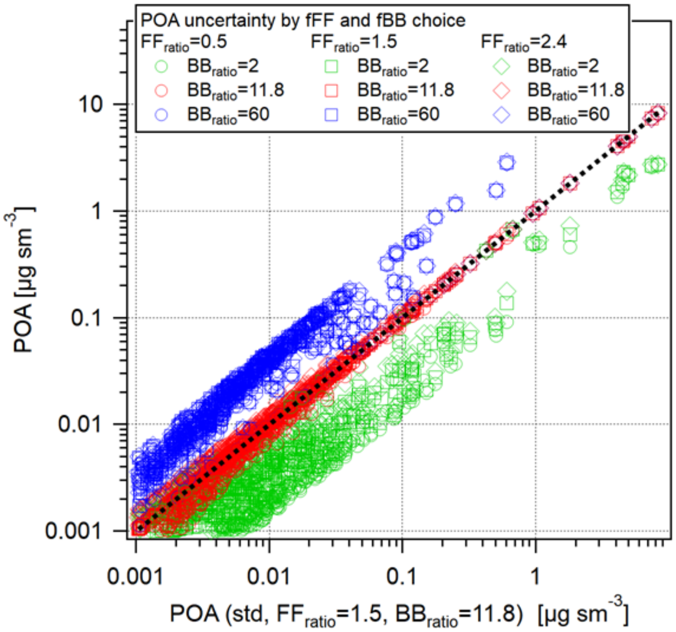
Hodzic presents a comparison of a large set of global models and observations from the free troposphere with a focus on organic aerosol. I applaud the authors on this large undertaking in terms of number of models and synthesis of observations. The manuscript contains a large amount of information. My major comments are regarding the POA estimation and better illustrating the utility of DYN configuration (which may just be a bit buried).

We thank the reviewer for the encouraging evaluation. We have modified the manuscript to address his/her concerns on the POA and DYN model configurations and other topics.

Major comments:

R3.1). Section 3.2 and POA estimation: Can the uncertainty in the POA estimates be quantified, ideally with error bars (e.g. in Fig 6 vertical profiles or Fig 8 POA)? Some ideas to consider regarding the POA estimation:

The uncertainty on the POA estimation is dominated by the choice of the biomass burning POA/BC emission ratio. We have documented both in the SI and in A2.5 at length that except for the pure BB points the uncertainty in the distribution of f(POA) is rather small (and for the purer BB points it's biased high toward more POA). Regarding OA, the following figure shows the uncertainty range from Figure S9 in concentration space:



So in absolute terms, the uncertainty in the estimation of total POA is about 20. But this does barely affect the actual fractions shown in Figure 6, since only a very small number of points actually contributes to the larger concentrations. For Figure 8 the uncertainty is likely larger, but will mostly affect the absolute numbers, not the trends with convection, which is the point of that figure.

R3.1-a) There are anthropogenic POA sources that do not have significant amounts of BC. See, for example, Figure 2 of Reff et al. (https://pubs.acs.org/doi/10.1021/es802930x) which indicates large emissions of dust associated with anthropogenic activity (road dust, construction dust) that have high OC relative to BC. While fossil fuel-BC may be a general proxy for anthropogenic activity, is the ratio 1.55 reflective of that general behavior? What does the EPA NEI indicate as FFratio including all anthropogenic PM sources (e.g. Fig 2 of Reff et al.)?

A3.1-a) We thank the reviewer for providing this additional reference. We have included values for the EPA NEI inventory in the supplementary materials (see Table S2). The associated ratios for the traffic and residential combustion sources in the NEI inventory (for traffic = 1.8, and residential = 8.2) are similar to those reported for other inventories and already discussed in the paper (section 3.2). Our ratios do not include emissions associated with fugitive dust from road, tire and construction, which is typically found in larger particles (D_{aero} > 1 µm, Zhao et al., 2017).

This is now explained in the manuscript:

"*It should be noted that urban model ratios do not include emissions associated with fugitive dust from road, tire and construction, as those are typically found in larger particles than those studied here (Zhao et al., 2017).*"

Zhao, G., Chen, Y., Hopke, P.K., Holsen, T.M., Dhaniyala, S.: Characteristics of traffic-induced fugitive dust from unpaved roads, Aerosol Science and Technology, 51:11, 1324-1331, DOI: 10.1080/02786826.2017.1347251, 2017.

Also see response to R3.1-d. Because we are determining the ratios from urban ambient air measurements, sources that do not emit BC (but do contribute on the submicron range) are also implicitly included.

R3.1-b) Can you plot data (obs or model) as a function of f(BB) influenced to get an intercept at f(BB)=0 and f(BB)=1 for comparison to FFratio and BBratio specified in text?

A3.1-b) Below are the requested quantile plots. On the left we are showing the ratio of the mean OA to mean BC quantiles, while on the right we are using medians. As already shown in Figure S11a and emphasized in A2.5, it highlights for the vast majority of cases along the ATom track, BC was a very small contributor to particulate carbon. And it also shows, as discussed in the manuscript, that the measured OA/BC in the BB plumes encountered during ATom were actually lower than the global average BB_{ratio} we use in our estimation. These plots have been added to the paper as Fig. S19.

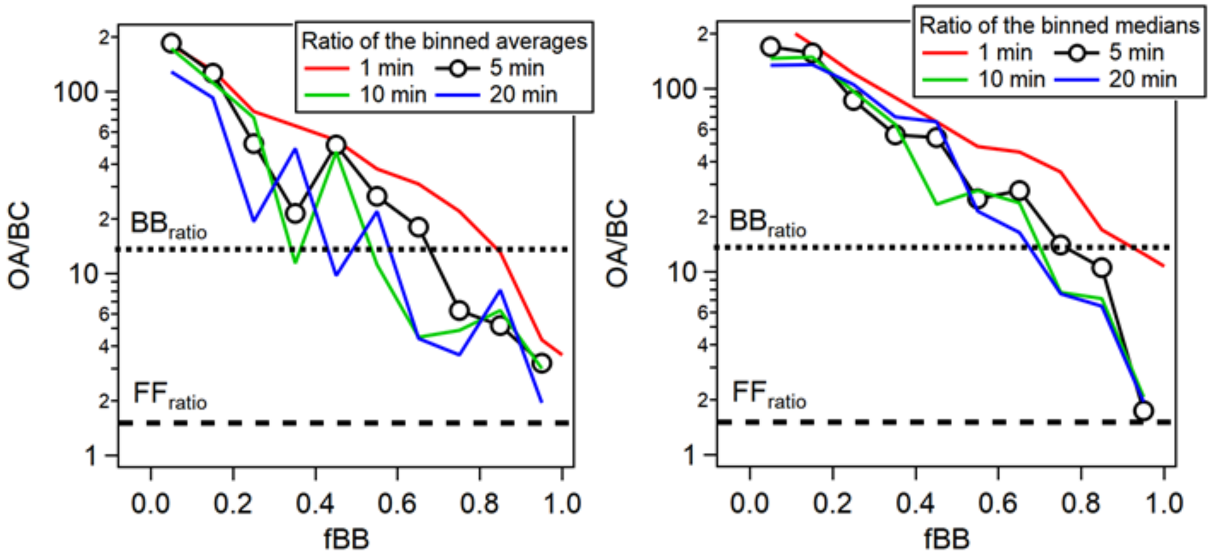
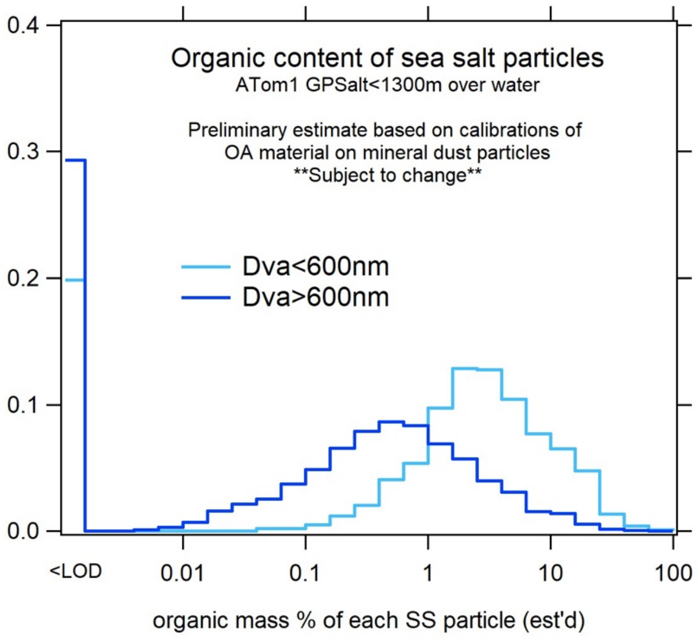


Figure S19: Distribution of the OA/BC ratio as a function of the fraction of BB influence measured by $f(BB)_{PALMS}$, calculated both as binned averages (left) and binned medians (right) for AToOm-1. Also shown are the OA/BC ratios that we currently assume based on the literature review for both anthropogenic (FF_{ratio}) and biomass burning sources (BB_{ratio}).

R3.1-c) For the calculation of observed POA, should there be small amounts of OA associated with sea spray and dust? You have sea salt and dust observations in Figure 2b and elsewhere that could be used to estimate that POA.

A3.1-c) We have added the following text to address this point:

"The contribution of POA from sea spray is difficult to constrain. As an order-of-magnitude estimate, marine POA is roughly calculated based on preliminary calibrations of OA on mineral dust particles from the PALMS instrument (personal communication K. Froyd). Using this calibration, the average OA by mass on sea salt was <10% for the large majority of MBL sampling (>85%). Since sea salt contributed 4% (11%) of mass in the AMS size range for ATom-1(2) (Figure 2), we estimate that marine POA is on the order of ~1% of aerosol mass in the AMS size range, and possibly much lower. Thus we think that it is reasonable to neglect the contribution of marine POA to this dataset. Future studies will refine this estimate."



POA associated with sea-salt particles in the marine boundary layer (<1300m) as reported by the PALMS instrument during ATom-1.

R3.1-d) Would it be better to label the estimate from equation 1 as "combustion POA"? I would assume models are also more specifically combustion POA?

A3.1-d) It would be inaccurate to label the estimated POA as combustion in the current model outputs as we cannot separate the combustion-emitted POA from other emission sources. This would require adding additional tracers and redoing the model simulations.

For the measurements, it would again be inaccurate to consider the estimated POA as only due to combustion. We have added the following text to the manuscript to clarify this point:

"The studies used to derive the emission ratio used ambient data in urban air, where all sources mix together and impact the POA/BC ratio, and thus the ratios include the impact of POA sources that may not emit BC."

R3.1-e) Is non-differential removal of BC and POA appropriate if POA is semivolatile? Consider that near a fire, concentrations could be high enough for IVOCs to be partitioned into the particle. As the airmass is diluted, POA will decrease more rapidly than a conserved tracer. This may explain why the larger BBratio gives POA >100% (line 378).

A3.1-e) This would be true if we were using the directly measured POA. But we are deriving POA from BC, and our method to estimate POA implicitly assumes that POA is non-volatile, and that it does not evaporate (and that also it is not lost to other processes such as photolysis or heterogeneous oxidation, even though those processes are known to be active in the atmosphere). This is done to obtain an upper limit for the measured POA, so that the key result that POA is too high in the models is reinforced. Therefore, POA > 100% in African plumes cannot be due to POA evaporation (and is instead likely due to uncertainties associated with the use of the global BB average ratios for those specific plumes; in fact the measured OA/BC value is within the combined uncertainty range for tropical forest given by Andreae (2019) 2.9..24.7).

The sensitivity of our results to the semi-volatile nature of POA has already been discussed in A2.3.

R3.1-f) For model estimates of POA, is hydrophilic OC considered POA or SOA? Can models just label the hydrophilic OC as "SOA" and get the right properties for endpoints (health, climate) of interest?

A3.1-f) Primary emitted hydrophilic OC is considered as POA. Indeed, the accepted definition is that for carbon emitted in the particle phase, even if it reacts in the particle phase, remains POA. For SOA to be produced from POA, POA needs to evaporate, and the evaporated organics gases to undergo gas-phase oxidation and then recondence into the particle phase. This is not happening in the model for the hydrophilic POA. This is consistent with earlier studies e.g. Pye and Seinfeld, 2010; Tsigaridis et al., 2014; Pai et al., 2019.

R3.1-g) Did you consider using a CTM to verify the method in equation 1 or back calculate the model effective FFratio? Does it reproduce the model POA (hydrophobic OC)?

A3.1-g) As already explained in the response A2.5, we did not apply this methodology to the model outputs as models do not track separately BC for various sources.

R3.1-h) BC is chemically aged. Do measurement techniques measure BC effectively at all atmospheric lifetimes?

A3.1-h) There is no evidence (that we are aware of, nor supplied by the reviewer) in the literature or in our measurements of refractory BC (such as measured in ATom) undergoing chemical loss to the gas-phase in the atmosphere. However, there are well-known aging effects (including coagulation, condensation, cloud processing) that do change the microphysical arrangement of BC by causing its physical shape to change and associating it with increasing amounts of internally mixed materials. Extensive testing as published in the literature (for example, Cross et al., (2010)) has shown that the measurement technique used in ATom to measure BC concentrations is insensitive to these aging effects. We have added the following text to the manuscript to address this point:

"Note that BC can physically age but it is not lost in any significant amount to the gas-phase due to chemical processes in the atmosphere."

Cross et al., (2010) Soot Particle Studies—Instrument Inter-Comparison—Project Overview, Aerosol Science and Technology, 44:8, 592-611, DOI: 10.1080/02786826.2010.482113.

R3.1-i) Line 650: The observation-method is not necessarily an upper-bound limit on the fraction of POA as it does not consider OA emitted in sea spray, dust, and may not consider all anthropogenic forms of POA. Consider rephrasing and/or demonstrating it is a limit by adding error bars by using more conservative (higher) FFratios.

A3.1-i) We respectfully disagree with the reviewer's statement. As explained in response A3.1-d above, our method implicitly accounts for all pollution sources of POA, even if they do not emit BC. As documented in response A3.1-c above, the estimated impact of marine POA in the submicron range during ATom is very small, and does not affect our results. Any amount of POA present in dust would be even smaller, as the dust concentrations in the submicron range (Fig. 2) are even smaller than those of sea salt, and the fraction of OA in dust is also very low.

R3.2) Line 355: Can you elaborate on why the PMF doesn't work to separate SOA and POA? When (what timescale, location, altitude, other factor) does the PMF stop working?

A3.2) Extracting linear factors by PMF always "works" in a technical sense, the question is whether it can provide the information that one seeks. This depends on the information content of the dataset and the questions asked. We have updated the text at this location to document this issue in more detail (since we do get this question relatively frequently), as:

"This approach is not suitable for ATom. To accurately resolve a minor factor such as POA in an AMS dataset, there needs to be a combination of: (a) Sufficient OA mass concentration, so that the signal-to-noise of the spectra is sufficient; (b) Enough fractional mass for the factor to be resolved (>5% in urban areas per Ulbrich et al. (2009), probably a larger fraction at low concentrations such as in ATom); (c) Sufficient spatio-temporal variability ("contrast") in the relative contributions of different factors, since that is part of what PMF uses to extract the factors; (d) Sufficient difference in the spectra of the different factors (for the same reason as (c)), and (e) relatively invariant spectra for each factor across the dataset (as this is a key assumption of the PMF algorithm). As an example of a near ideal case, in Hodshire et al (2019) we extracted MSA by PMF from the ATom-1 data, and were able to match that factor with our independently calibrated MSA species. A very distinct and nearly invariant mass spectrum was measured repeatedly near sources (MBL) (and was mostly absent elsewhere, thus providing strong spatio-temporal contrast) and accounted for about 6% of the fractional mass and 15% of the variance in time. Thus all the conditions were met. For POA, on the other hand, the air sampled in ATom and

coming from e.g. Asia has POA and SOA very well mixed, with little change on their relative mass fractions vs. time (as the aircraft flies through that airmass). POA is very low, as documented later in this paper. Atmospheric aging makes the spectra from all OA sources more and more similar as measured by AMS spectra (Jimenez et al., 2009). Thus most of the conditions above are not satisfied for extracting POA by PMF analysis of this dataset."

Ulbrich, I.M., M.R. Canagaratna, Q. Zhang, D.R. Worsnop, and J.L. Jimenez. Interpretation of Organic Components from Positive Matrix Factorization of Aerosol Mass Spectrometric Data. Atmospheric Chemistry and Physics , 9, 2891-2918, 2009.

Hodshire, A.L., P. Campuzano-Jost, J.K. Kodros, B. Croft, B.A. Nault, J.C. Schroder, J.L. Jimenez, J.R. Pierce. The potential role of methanesulfonic acid (MSA) in aerosol formation and growth and the associated radiative forcings. Atmos. Chem. Phys., 19, 3137-3160, 2019, https://doi.org/10.5194/acp-19-3137-2019

Jimenez, J.L., M.R. Canagaratna, N.M. Donahue, A.S.H. Prevot, Q. Zhang, J.H. Kroll, P.F. DeCarlo, J.D. Allan, H. Coe, N.L. Ng, A.C. Aiken, K.D. Docherty, I.M. Ulbrich, A.P. Grieshop, A.L. Robinson, J. Duplissy, J. D. Smith, K.R. Wilson, V.A. Lanz, C. Hueglin, Y.L. Sun, J. Tian, A. Laaksonen, T. Raatikainen, J. Rautiainen, P. Vaattovaara, M. Ehn, M. Kulmala, J.M. Tomlinson, D.R. Collins, M.J. Cubison , E.J. Dunlea, J.A. Huffman, T.B. Onasch, M.R. Alfarra, P.I. Williams, K. Bower, Y. Kondo, J. Schneider, F. Drewnick, S. Borrmann, S. Weimer, K. Demerjian, D. Salcedo, L. Cottrell, R. Griffin, A. Takami, T. Miyoshi, S. Hatakeyama, A. Shimono, J.Y Sun, Y.M. Zhang, K. Dzepina, J.R. Kimmel, D. Sueper, J.T. Jayne, S.C. Herndon, A.M. Trimborn, L.R. Williams, E.C. Wood, C.E. Kolb, A.M. Middlebrook, U. Baltensperger, and D.R. Worsnop. Evolution of Organic Aerosols in the Atmosphere. Science, 326, 1525-1529, 2009. doi: 10.1126/science.1180353.

R3.3) Perhaps the strengths of DYN could be better isolated/highlighted (e.g. what life-times, OA regimes, POA levels, etc does DYN perform better for?). The abstract statement: "concept of a more dynamic OA. . .with enhanced removal of primary OA, and a stronger production of secondary OA in global models needed to provide a better agreement with observations" could use more support.
Evaluation seems mixed: Figure 8, indicates GC12 and CESM1-CARMA have reasonable POA (while CESM2-DYN has overestimated POA). Line 803 indicates DYN evaluates better, but that effect seems marginal or secondary to other issues in Figure 9 b at least for GC. RMSE in Table 2 also tends to increase when going from base to DYN treatment in CESM2 and GC12. Figure 7 improvements seem mixed.

A3.3) We agree with the reviewer that given the extreme complexity of the dataset and the models, it is difficult to identify the model configuration that works the best all the time. We do not conclude that DYN is providing better results than other models under all conditions (e.g. CESM2-DYN still has a problem with the POA removal in convection as the fix implemented in CESM-CARMA has not yet been implemented in this version). We have reworded the text in section 4.1 to better describe the model behavior in the NH summer and winter, and indicate more clearly when and where the stronger SOA production provides better results:

"*During the NH summer (ATom-1), models using the VBS parameterization from Pye et al. (2010) tend to underpredict the OA concentrations by 43% for GC12-REF and 33% for CESM1-CARMA for ATom-1, most likely due to the excessive evaporation of the formed SOA in remote regions and low yields for anthropogenic SOA (Schroder et al., 2018; Shah et al., 2019). Models using the VBS parameterization from Hodzic et al. (2016) (CESM2-DYN and GC12-DYN) where OA is less volatile and also OA yields are corrected for wall losses show an improved agreement with observations especially for CESM2-DYN (with NMB of ~5%), and to a lesser extent for GC12-DYN (NMB of ~33%). During the NH winter (ATom-2) characterized by a lower production of SOA, both VBS approaches lead to an overestimation of the predicted OA. This is likely caused by excessively high levels of primary emitted OA as discussed in section 4.4.*"

More importantly (not specific to DYN), in this study we have shown as stated in the conclusions (and abstract) that "*the OA system seems to be more dynamic with a need for an enhanced removal of primary OA, and a stronger production of secondary OA in global models to provide a better agreement with observations.*" And that is supported by the fact that models that have improved in-cloud removal of POA tend to perform better with regard to POA concentrations and vertical profiles in the upper troposphere than those that don't have it (e.g. CESM1-CARMA with improved in-cloud removal vs. without improved in-cloud removal in Figure 8 and S6; GC12 in Figure 8 and S6 compared to CESM2). We have also shown that making POA semi-volatile instead of non-volatile in model simulations aggravates the model bias in the upper troposphere, and that removal by deep convective clouds and possibly photolysis are needed to address model bias.

The need for a stronger SOA production in models is supported by the fact that models that have correct POA (CESM1-CARMA or GC12) need a stronger production of SOA to match the measured concentrations. The comparison of GC12-REF with two SOA formation mechanisms (using the same removals) in Figure 8 illustrates that a stronger production of SOA would lead to an improved agreement with measurements.

Minor comments:

R3.4) Abstract states "OA predictions from AeroCom Phase II...span a factor of 400- 1000. . ." should that be the inter-model variability spans a factor of 400-1000 or the concentrations predicted span that range?

A3.4) The first one. This has already been clarified in response A2.1 above.

R3.5) Did all models assume nonvolatile POA?

A3.5) Yes, see response A2.3.

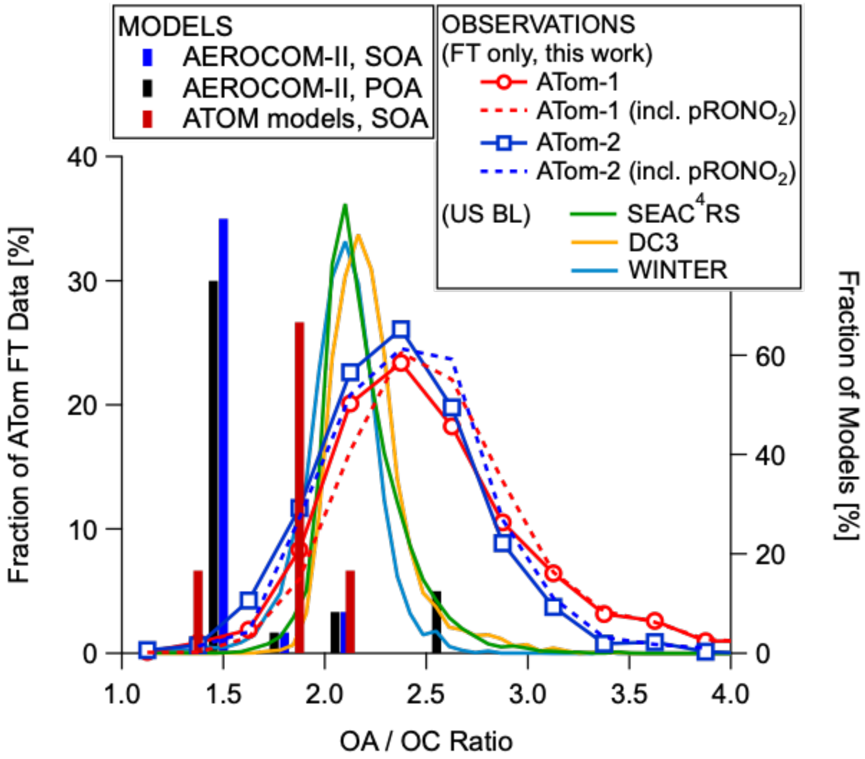
R3.6) Do OA/OC ratios include consideration of S and N in the form of organosulfates and organonitrates?

A3.6) The following text has been added to the manuscript to address this point:

"Note that for organosulfates (R-O-SO_2H and organonitrates (R-O-NO_2, p$RONO_2$ in the following) only one oxygen is included in the reported OA/OC, as the fragments of these species are typically the same as for inorganic species in the AMS (Farmer et al., 2010). However in ATom organosulfates are estimated to account for ~1% of the total sulfate (based on PALMS data, see Liao et al., 2015 for the methodology). Since sulfate and OA concentrations are comparable, organosulfates would only increase the OA/OC by ~1% on average. Organonitrates are reported from the AMS for ATom. Their impact on OA/OC is not propagated for the default values, to maintain consistency with a large set of OA/OC measurements by AMS in the literature, and since they would increase OA/OC on average by only 4.5% (ATom-1) and 2.2% (ATom-2), which is smaller than the uncertainty of this measurement. However, we show the results with both methods in Fig. 5 to fully document this topic.

Farmer, D.K., A. Matsunaga, K.S. Docherty, J.D. Surratt, J.H. Seinfeld, P.J. Ziemann, and J.L. Jimenez. Response of an Aerosol Mass Spectrometer to Organonitrates and Organosulfates and implications for Atmospheric Chemistry. Proceedings of the National Academy of Sciences of the USA, 107, 6670-6675, doi: 10.1073/pnas.0912340107, 2010.

Liao, J., K.D. Froyd, D.M. Murphy, F.N. Keutsch, G. Yu, P.O.Wennberg, J.St. Clair, J.D. Crounse, A. Wisthaler, T. Mikoviny, T.B. Ryerson, I.B. Pollack, J. Peischl, J.L. Jimenez, P. Campuzano Jost, D.A. Day, B.E. Anderson, L.D. Ziemba, D.R. Blake, S. Meinardi, G. Diskin. Airborne organosulfates measurements over the continental US. Journal of Geophysical Research-Atmospheres, 120, 2990–3005, doi:10.1002/2014JD022378, 2015.



***Caption Fig. 5**: Distribution of the OA / OC ratio as measured during ATom-1 and -2. Also included (in dashed lines) are the distributions of OA/OC values that included the contribution of organic nitrates (pRONO$_2$). Values for the recent aircraft campaigns (SEAC4RS, DC3 and WINTER) that took place over continental US regions closer to continental source regions are also shown (Schroder et al., 2018). The bars (right axis) show the OA/OC used for SOA and POA by the models included in the AeroCom and ATom ensemble, with OA/OC=1.4 being the modal value for the former and 1.8 for the latter.*

R3.7) At times, it was a bit confusing if the AeroCom-II ensemble referred to circa 2014 models or the actual AeroCom II results paired with measurements. On line 103, the “AeroCom results” could be clarified as “AeroCom models” or “models used in AeroCom.” Line 277 might be better as “monthly average results of 28 global models” instead of “results of 28 global models.”

A3.7) We have modified the text as suggested by the reviewer: “*Our own analyses of the AeroCom-II models..*” and “*We consider the monthly average results of 28 global models,..*”.

R3.8) Figure 1: Is “author affiliations” supposed to be in the caption?

A3.8) We have added *“Model contributor affiliations”* to the figure label and caption. See also response A1.3.

R3.9) Line 163: While marine production of OA may be smaller than continental production, it's possible marine production contributes more to concentrations over the oceans than it's global production rate would suggest. I recommend rewording sentence. What concentrations of SOA are predicted for oceanic isoprene sources?

A3.9) SOA production from oceanic isoprene sources is not included as already indicated in the manuscript.

As suggested by the reviewer, we have modified the sentence to read: "*None of the models includes the marine production of OA which is estimated to be ~3 orders of magnitude smaller than the continental production of OA from both isoprene and monoterpene precursors (Kim et al., 2017), but could be important in the MBL.*"

See also the response to A3.1-c above.

R3.10) How does averaging data allow below detection values to be used?

A3.10) This seems to be a very widespread issue with modelers' understanding of our measurements. We have included the following text to section 3.3. to clarify and document this point:

"Per standard statistics, the precision of a measurement decreases (i.e., gets better) with the square root of the number of points (or time interval) sampled. I.e. the precision of an average can be approximated by the standard error of the mean (σ/sqrt(n), where n is the number of measurements averaged), and it is better than the precision of the individual data points (σ). This also applies to the detection limit, since it is just 3 times the precision. Note that a detection limit is not meaningful unless the averaging time is specified. For example, let's assume that the detection limit is 20 ng m^{-3} (1-second), and the data points over 60 consecutive seconds are all 10 ng m^{-3}. All 1-second measurements are below the 1-second DL. However the average (10 ng m^{-3}) is now above the DL for 1-minute averages, which is 20/sqrt(60) = 2.6 ng m^{-3}."

R3.11) Line 605: qualify the OA/OC of 1.4 as "fossil fuel" combustion as biomass combustion tends to have OA/OC of 1.7.

A3.11) We have added this qualifier as requested: "*A low OA/OC ratio is indicative of freshly emitted OA from fossil fuel combustion (typically ~1.4),..*"