29 march 2023

A second review led to requests for additional clarifications. The reviewer comments are shown in black, with the author responses shown in blue and any edited manuscript language shown in italicized blue font.

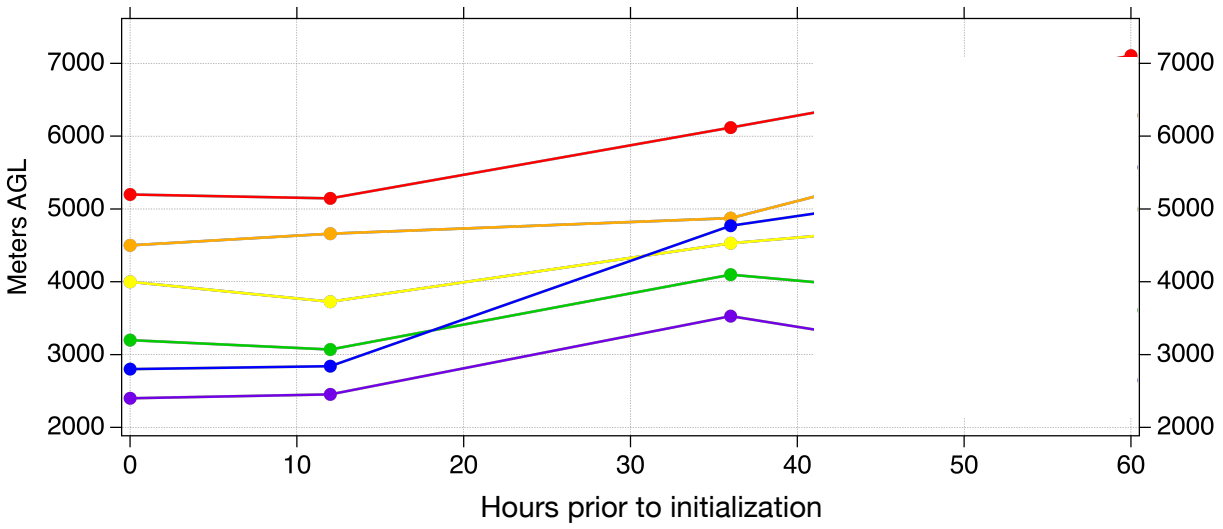
The reviewer recommends the manuscript for publication after considering the following points. The uploaded manuscript reflects the updated language contained herein.

Line 115: Are trajectories shown?

Backtrajectories are shown in Fig. 10. We have changed the sentence on line 115 to make clear that we show them ("*Backtrajectories based on the HYSPLIT model \citep{Stein15}, also driven by the NCEP GFS meteorology, further illuminate the pathway the BBA traveled prior to its sampling on 24 September 2016 (shown later in Fig. 10).*").

Lines 116-118: Why is the time for vertical transport not accounted for? Can you explain.

The NCEP back trajectories rely on the GFS large-scale vertical velocity. Their altitude change over the three days prior indicates the smoke has descended to where it was detected by the plane, shown below for the trajectories in Fig. 10. We do not account for the time that it takes the smoke to travel from the surface to altitudes of 3.5-6km. Fig. S3, based on the time estimate from model-released tracers, suggests it might take about 3 days for the smoke to reach the higher altitudes above the surface fire emission sources, but this is likely a model-dependent value. WRF-AAM places all the smoke in the surface layer and then lets the model advect the smoke upwards.

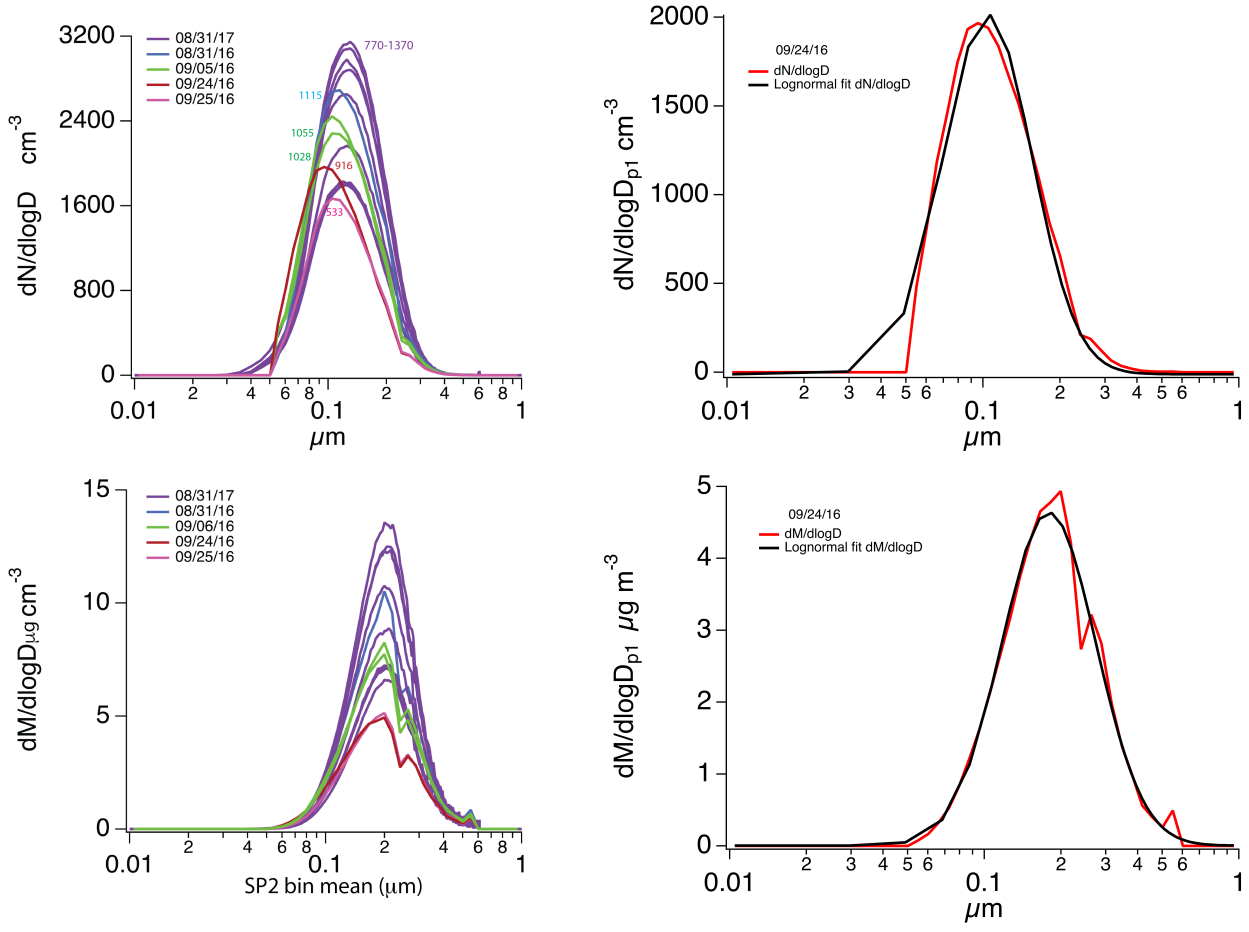


We have expanded the language around lines 116-118 to now read as: "*These backtrajectories reach the location of the fire emission source after approximately three days, at higher altitudes than where the aerosol was sampled. The aerosol age estimate is younger than the WRF-AAM aerosol forecasts would suggest. This is because the time needed for the aerosol to travel from*

the surface to the higher altitude is not accounted for in the HYSPLIT age estimate. The vertical advection timescale is highly model-dependent on boundary layer parameterizations and we merely make note of the difference in the two age estimates here."
Is this clear enough?

Line 138: A detection of 99% of the BC mass is relatively high. Does this value include results from fitting the BC size distribution taking into account the calibrated SP2 detection efficiency? Is this value the same for all considered periods.

This and a latter comment by the same reviewer motivated us to dig out the full SP2 size distribution data (the campaign archived dataset only contained the total SP2 number and mass concentrations). We thank the reviewer for pushing us to do this. We now show the full number and mass distributions for the level legs detailed in Table S2 as the new figure 3, shown to the left below. Log-normal fits (examples shown in the new Fig. S4 and included here below right) indicate the upper limit of the SP2 size range - 0.5 micron - captures the upper end of the size range. At the lower end, the drop off in detection sensitivity for diameters < 0.08 micron is visible in the number size distributions for the 8/31/2016, 9/24/2016 and 9/25/2016 sample. Also evident though is that the contribution to the total mass at these small sizes is small - close if not quite at the 99% number reported within Taylor et al., 2020 - and can basically be neglected. Because of the instrument artifacts visible in the mass size distributions below, the difference in the measured mass and mass calculated from the log-normal fit was not dominated by the drop-off in detection efficiency for the smaller sizes and we do not report those numbers - but they are close to each other.



We have rewritten the language for this section as: "*Size distributions by number and mass (the latter assuming a BC density of 1.8 g cm^{-3} \citep{Bond06}) for the level legs detailed in Table S2 indicate particle diameters that are well below the upper SP2 detection limit (Fig.~\ref{rBC_dist}; the kink at 2.5 μm for the 2016 samples is attributed to an instrument artifact). Lognormal fits help visualize a drop-off in detection efficiency for diameters < 0.08 micron for the samples weighted towards the smaller sizes (Fig. S4, for 9/24/2016 and 9/25/2016). The SP2 size range nevertheless captures almost all of the black carbon mass, close to the 99\% value reported for CLARIFY \citep{Taylor20}. rBC total number concentrations vary between 530 to 1370 cm^{-3} (Fig.~\ref{rBC_dist}a), and undercounting of the mass and number through coincidence is estimated to be less than 3\% based on \cite{Taylor20}.* "

Line 162: The reviewer thinks that the lower plot in Fig. S4 is important enough to be shown in the main text, because the selection of the analyzed periods (OA>20μg/m3) is based on OC:BC versus OA.

Fig. S4b is now Fig. 4.

Line 165: The authors refer to figures S11 and S13. It is difficult to estimate how the threshold of 10 μg/m3 affects the results without seeing the same figures for 20 μg/m3 next to it.

These figures can be compared next to each other, either on a screen or as a printout. The trends (or lack of) are the same in OA:BC as a function of f44, the chemical aging marker, for both thresholds. We did find a slight difference in BC:CO between the two thresholds for the data from 8/31/2016 and the stricter threshold led us to reject the 8/31/2016 from consideration (we had originally hoped to include it but there also wasn't enough data available with the stricter threshold). We primarily include the figures in response to an earlier reviewer. We did change the language here to say: "*The BC:ΔCO and OA:BC ratios are shown for individual flights as a function of {\it f44} for both thresholds in Figs. S10-S13 in the interest of full documentation.* "

Line 189: SEA is not defined.

Thanks for catching, fixed.

Line 205: The authors arguing with the BC size distribution to explain differences between LDMA and UHSAS. Is there are reference to BC size distributions?

We show the BC size distributions now in Fig. 3. These confirm that the larger particles detected by the LDMA/UHSAS, in the 0.15-0.2 micron range, are more likely to be the ones containing the black carbon. We have added a sentence here: "*Two particle populations emerge: one with diameters between 0.15-0.2 μm that is more likely to contain black carbon, (see Fig.~\ref{rBC_dist}), and another with median diameters $<$ 0.1 μm, speculated to consist mostly of OA.*".

Line 213: Uncertainty of 30% in number or volume?

Number. Now mentioned in the text.

Line 215: Should it be 'The black carbon core mass-median diameter is also estimated using the SP2-provided mass size distribution …' ?

Yes. Now that we have the full SP2 size distribution we can say this. Since we emphasize the SP2 size distribution earlier, we now just state "*The black carbon core mass-median diameter is used to infer fire conditions at the source*".

Line 229: Specify where it can be found in the supplement.

We now mention '*Section 7 of the Supplement*' instead of 'Supplement' on line 229.

Line 311: The reviewer believes that Fig. 7a does not support the conclusion that MAC is correlated to FrBC.

We have rewritten the sentence as '*The bulk mass absorption coefficients (MAC$_{BC,660}$) decrease with the estimate of the fraction of particles containing black carbon (FrBC) for the ORACLES-2016 flights (Fig. 7a), with the SSA values decreasing more systematically with FrBC for all four flights (Fig. 7b).*'

Line 337: How was TC measured?

TC is calculated as TC=BC+organic carbon (OC) and OC is estimated from OA:OC=1.26*O:C+1.18 (Aiken et al., 2008), same as is done within Brown et al., 2021. We included this information within the caption of Fig. 8. We have moved this information to the main text based on this comment.

Line 410: Is there any evidence in the literature that a significant mass loss by this mechanism is possible?

The reviewer is pointing to mass loss by oxidation through fragmentation. Jimenez et al., 2009 and O'Brien and Kroll, 2019, both cited, were influential in our thinking. We are only speculating -- and certainly if there is a better idea we would love to hear it. We do not detect much brown carbon and are looking for explanations that do not involve brown carbon. Fig. 12 makes clear that mass loss is genuinely occurring

Figure S1: Color scales not correct or overlap with other scales

I think the reviewer is referring to the top left hand figure, for which the total aerosol optical thickness color scale differs from that of the other aerosol forecasts. This was at the very beginning of the ORACLES-2016 campaign and the forecasts were still under development. This image is all we have.

Supplement line 115: The Anderson and Ogren (1998) correction is used only to correct the nephelometer.

Correct. But this is already stated in the supplement.

Supplement line 116: Heating to 50 °C can lead to a loss of volatile material. This could reduce the light enhancement factor by making the coating thinner. In addition, the mass loss could affect the correction of the light scattering artifact when using the Virkkula (2010) correction. Does this significantly affect the derived light absorption coefficient and the single scattering albedo?

As shown in Fig. S6, the size TDMA size distributions differ little between unheated samples and samples heated to 150C. The lack of volatility is consistent with the OA regimes shown in Fig. 6 (the ORACLES aerosol are primarily classified as ‘aged’ and ‘semi-volatile’ regimes). Barrett et al., 2022 (cited) compared the ORACLES PSAP absorption values to the more accurate absorption measurements by the CLARIFY EXSCALABAR and DOE CAPS-SSA instrument, and these corresponded well. Based on these various independent assessments we claim that any loss of additional volatile material by heating to 50C over a short time period will fall within the measurement uncertainties.