

Atmos. Meas. Tech. Discuss., author comment AC1
<https://doi.org/10.5194/amt-2022-59-AC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

Paul A. Barrett et al.

Author comment on "Intercomparison of airborne and surface-based measurements during the CLARIFY, ORACLES and LASIC field experiments" by Paul A. Barrett et al., Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2022-59-AC1>, 2022

Response to reviewers: Intercomparison of airborne and surface-based measurements during the CLARIFY, ORACLES and LASIC field experiments

Paul Barrett et al.

August 2022

We thank both reviewers for taking the time to read through this paper and offer many constructive criticisms that have no doubt improved the manuscript. We recognise that the manuscript is long and the results were not presented as concisely as they could have been. We have attempted to rectify this through use of additional figures and removal of some tabulated materials to the Supplement. Whilst the text could have been shortened with the use of tabulated information about the instrumentation we felt that readability would have suffered and so kept the section broadly the same.

Use of ODR fitting was undertaken initially but we have taken onboard the suggestions to shorten the averaging period and have done so where possible and have now also included uncertainties on the ODR fit parameters. We now described the method in detail at the head of the results section.

We have concentrated on primary measured quantities and so moved some derived parameters such as dew point temperature, relative humidity and aerosol particle effective radius to the supplementary materials.

Some of the discussion has been moved in to the results section, including that around thermodynamics to make manuscript more readable. The discussion section is now more focussed on synthesis of results and outstanding issues, such as the impact of inlets, etc.

We present the responses to both Reviewer #1 and Reviewer #2 below. Comments are copied in grey italics for convenience. We do not include every change to the manuscript in here as that would be unwieldy, so we also upload a marked-up manuscript with differences highlighted. We have added references to relevant literature that has become available since submission.

1. Responses to Reviewer #1 1.1 Major

1) *The comparisons between the various instruments are based primarily on linear regression against mean values from long periods of flight. There are several problems with this approach:*

- *a) The uncertainties quoted are for each instrument's inherent response time as installed*

in the aircraft. Yet averaging together many minutes of data will result in reduced uncertainties (if the same population is being randomly sampled). One would expect better agreement than the stated raw instrument uncertainties for such averaged data.

- *b) Regression should be applied using the highest time resolution data possible, rather than to just a few average values from these different "runs". A quick example: if there were only two "runs", using this manuscript's approach there would be only two values, and regression would be a perfect fit to the two data points. The agreement between instruments should be based on the highest resolution data reported, to which the stated uncertainties apply. If one were to fit to averaged values, uncertainties must be adjusted and accounted for in the regression. It would be very interesting to see the regression from the large dynamic range covered in the profile of the two aircraft; this would be a nice way to rigorously compare instruments in a challenging environment.*

- *Fits were performed using ODR originally, but this was not stated explicitly. Regressions have now been redone - and performed on 10s segments rather than flight leg averages. See below for details.*

1 a), b)

The datasets tend not to be valid at the raw instrumental resolution due to the nature of sampling from the different platforms. In particular due to sampling through inlet systems and through pipework which can result in physical "smoothing" of the signals due to imperfect transport, and possible temporal offsets – which whilst we attempt to correct for this may still be present. Small timing errors may differ between instruments on the same platform and between platforms. In most instances e.g. optical absorption and extinction on FAAM the true fastest response possible has been demonstrated in the laboratory to be between 6 and 10 seconds. Therefore, we have first smoothed data to 10 s (i.e. 0.1 Hz) from the aircraft.

Data have been included for as wide a dynamic range as possible from the full flight intercomparison section. This includes the very dry and relatively clean troposphere at close to 6km and the polluted humid oceanic boundary layer. We do not include data specifically from the whole profile, as many instruments are not optimised for use during descents / pressure changes. We now show the data from the absorption measurements, and the problems can be seen in the artefacts in the NASA PSAP data, where there is a spike in data on red and blue channels, resulting in unrealistic looking single scattering

albedo values. We feel that using the data from known good times in the free troposphere leg and the descent through the pollution layer in the free troposphere is a good compromise. We have also used observed CLARIFY PAS observations data to compute Angstrom exponents for all wavelength pairs for the airborne comparisons, rather than relying on the campaign mean from Taylor et al. (2020) as we had done originally.

Concentrations of pollutants, chemical and physical varied over the range that is presented – we do not include data that is below demonstrated (in the laboratory) limits of detection.

Data from LASIC must be treated differently, as the measurements are offset in space and time. Here we keep the observations as mean values and variability.

The errors in x and y and the ODR fits are taken as the standard error over the averaging period. We have now added commentary at the start of the results section that gives details of the method and the reasons for the choices made in the analysis. We are aiming to find the similarity or differences between the observations on two platforms, rather than construct a function that maps one set of observations on to the other. Of course, should downstream users want to obtain measurements with reduced uncertainties then they could average over any length of time of their choosing, considering natural spatial and temporal variability and we expect them to do this on a per-instrument basis as they require.

The fit parameters only changed by minimal amounts (a few percent), by changing from run averages to 10 s data - for example:

	Original ODR fit	New ODR fit
CO	$8 + 0.97x$	$9.5 + 0.95x$
O3	$-1 + 1.19x$	$-9.6 + 1.17x$
σ_{SP} at 660 nm (PM10)	$-0.1 + 1.56x$	$-0.57 + 1.52x$
σ_{SP} at 660 nm (PM1)	$-0.3 + 0.90x$	$-0.72 + 0.97x$

- c) *The linear regressions appear to use one-sided least-squares fits. Because there are*

uncertainties in both x and y parameters, a 2-sided regression, such as orthogonal distance regression, should be used to determine slopes and intercepts. Further, the regressions should account for the uncertainties in each parameter, whether averaged or not.

Fits are performed using orthogonal distance regression, this was not stated in the original manuscript.

2) Most of the data are presented in Table 3, which is so large as to be completely unwieldy and is extraordinarily difficult to read because it spans multiple pages. Generally it is much preferable to show data graphically. Instead of Table 3, I recommend a series of graphs of the key data that are discussed and analyzed (at their native resolution). For example, a plot of extinction coefficient for the two airborne platforms could be shown with all of the data covering the full dynamic range, with points perhaps colored by the run type (BL, FT, etc.). It may be most effective to use log-log plots to show the range of values clearly. The numerical values in Table 3 could go into an appendix or the supplemental materials, hopefully in a more compact format.

We agree that the table was too large.

New Fig. 5 now contains comparison plots of temperature and humidity, with the data from this portion of the table moved to the supplement. Aerosol number correlation plots are added to the PSD figure (new Figure 7). We have chosen log plots for humidity data and kept linear for others which we deem best to show the data.

Where possible we now show data points coloured by altitude. For some parameters we do not do this in order to preserve clarity.

Data for (new) Figs 5, 6, and 7 are shown as the 10 s values rather than run averages for airborne data.

Where data are now plotted the values from Table 3 are moved to the supplement. For the parameters that remain – they have been split in to sub-tables, and placed on landscape pages, reducing the number of pages of tables in the main manuscript. We have retained chemical composition measurements and derived properties as these are present only from the boundary layer at one point in time. The LASIC data (which compare badly) are included for completeness (item #5). Cloud physical properties are also tabulated as only one run was performed in cloud.

We are showing all the extinction data that it is possible to show – given the times we know that instruments were operating outside their valid operating parameters. For example, NASA extinction data requires scattering and absorption, but the PSAP which measures absorption does not perform well during the descent.

3) There is extensive discussion of aerosol number concentration and effective radius.

However, aerosol mass is extremely important as it is the parameter most often carried in models. Thus it would be very useful to compare integrated volume from the different size distribution instruments. I would suggest that Fig. 6 be converted to 6 panels, with a, b, and c showing, on a linear y-scale, the number concentration comparisons, and panel d, e,

and f showing the volume concentrations on a linear panel. A log-log scale with almost 9 orders of magnitude on the y -axis can hide so much detail. For example, at ~ 2 nm in the current Fig. 6a, there is almost an order of magnitude difference between the green line (FAAM PCASP1) and the others. Is this significant? When plotted on a linear scale we can see if this difference is a significant contributor to parameters we care about, such as integrated number or volume (mass).

3) We have modified the particle size distributions (new Fig. 7) to show number and volume distributions. Linear y -scales are used for both. We chose to keep the elevated pollution plume and free troposphere data on the same figures (b) and (d) as the purpose is to show that the instrument can differentiate between the weak pollution plume and the cleaner surroundings – at least for particle number distributions. The particle volume distributions are shown to be poor – as there are so few particles at the larger diameters. We do not show cumulative distributions because there is no good way to integrate number or volume across multiple probes, without creating a composite fit – and that is beyond the scope of this study and left for individual research questions.

4) Figure 8. I had trouble understanding Fig. 8b. The y -label says it is the Angstrom

exponent of absorption, but the caption says it is that for extinction. Is it derived using Eq. 2 applied to the absorption coefficient values shown in Fig. 8a? If so, why are the markers in 8b plotted at ~ 460 nm when the closest wavelength pairs are at 470 and 405 nm? Please explain carefully how these values were derived. Also, it would make more sense graphically for these two plots to be side-by-side, to enhance the vertical scaling and make differences more evident.

Corrected extinction to absorption in caption.

Wavelength pairs for blue, green absorption from FAAM EXSCALABAR are 405 nm and 515 nm giving a mean of 460 nm. NASA PSAP instrument has wavelengths 470 nm and 530 nm giving a mean of 500 nm, as plotted.

We have replotted the figure with one panel above the other as suggested and narrowed the aspect ratio to permit printing in a single column rather than spanning both.

5) Lines 950-956. *The agreement between the AMS on the FAAM aircraft and the ACMS at the ARM site was quite poor, with factors of 3-4.5 difference. These data should be shown in Table 3, but are not. Poorly agreeing data can be just as important as data that agree well, so please show the values if they are part of a project data archive and not rejected for quality-controlled reasons independent of this comparison.*

This data is included in the revised version (see item 2.)

1.2 Minor

1) *Abstract. The data are described multiple times as agreeing "well". This should be changed to a more quantitative statement, such as "the data agreed within combined experimental uncertainty", if this is the case when the comparison is made at time resolutions for which the stated uncertainties are valid (see comment 1b above).*

Abstract – modified to remove "well" and added context.

2) *Line 186. Need period after "Beer's Law"*

Added period.

3) Line 217. Two periods.

Removed period.

4) line 249. Change "dependant" to "dependent", here and elsewhere.

Changed dependant to dependent globally.

5) Line 255. I don't understand this sentence. Please clarify.

Removed the sentence.

6) Line 268. Do "rear" and "front" instruments refer to PSAPs or nephelometers?

Added PSAP to line 271 for clarification.

7) Line 283. Please state the flow rates to each optical instrument.

Added line 285 - "The nephelometer drew at 30 L min⁻¹ and the PSAP 2 L min⁻¹."

8) Line 379. What are representative uncertainties for the absorption coefficient

determined from the CAPS PMSSA instrument?

Added line 377 - "The CAPS PM_{SSA} measurement uncertainties for absorption coefficients are estimated in Onasch et al. (2015). For a typical SSA ~0.8 during LASIC, a conservative uncertainty estimate for the absorption coefficient is ~20%."

9) Line 397. Moore et al. (2021) provide a thorough analysis of refractive index

sensitivities for the UHSAS. Moore, R. H., Wiggins, E. B., Ahern, A. T., Zimmerman, S., Montgomery, L., Campuzano Jost, P., Robinson, C. E., Ziemba, L. D., Winstead, E. L., Anderson, B. E., Brock, C. A., Brown, M. D., Chen, G., Crosbie, E. C., Guo, H., Jimenez, J. L., Jordan, C. E., Lyu, M., Nault, B. A., Rothfuss, N. E., Sanchez, K. J., Schueneman, M., Shingler, T. J., Shook, M.A., Thornhill, K. L., Wagner, N. L., and Wang, J.: Sizing response of the Ultra-HighSensitivity Aerosol Spectrometer (UHSAS) and Laser Aerosol Spectrometer (LAS) to changes in submicron aerosol composition and refractive index, *Atmos. Meas. Tech.*, 14, 4517–4542, <https://doi.org/10.5194/amt-14-4517-2021>, 2021.

Added line 405 – "Moore et al. (2021) noticed similar behaviour in laboratory tests of a UHSAS for highly absorbing aerosols. Here we use the NASA P3 data for comparison with the outboard FAAM BAe-146 PCASPs." and reference to Moore et al. (2021).

10) Line 393. Although this is described in more detail in Wu et al. (2020), please provide a succinct explanation for why an empirical correction factor is needed for the SMPS, when it's quite a fundamental instrument.

Added line 393-399 – "Previously a comparison was made for CLARIFY data between estimated volume concentrations derived from AMS + SP2 total mass concentrations and PM1 volume concentrations from PCASP (assuming spherical particles). Estimated AMS+SP2 volumes were approximately 80 % of the PCASP derived values, which was considered reasonable within the uncertainty in the volume calculations (Wu et al., 2020) demonstrating consistency between inboard and outboard measurements. Discrepancies between SMPS (inboard) and PCASP (outboard) number concentrations remained however

and so the SMPS concentrations were reduced by a collection efficiency factor of 1.8 to give better correspondence in the overlap region of the PSDs. The cause remains unknown."

11) Line 403. Perhaps just state "with updated electronics" rather than "with SPP200 electronics". Or explain what SPP200 means.

Modified line 413 – "FAAM and NASA flew wing-mounted DMT PCASPs (Lui et al., 1992) with updated electronics (nominally SPP200, DMT 2021) which were exposed to the free airstream."

12) Line 417. Change "bin dimensions" to "bin boundary diameters".

Replaced bin dimensions with bin boundary diameters.

13) Line 418. The underwing PCASP is not only not adjusted for the "absorbing characteristics" of the BBA, but it's in general not adjusted for any varying refractive index, including water. This could make a significant sizing difference with in-cabin spectrometers.

This is discussed in results – new line 738-743 - Data for runBL were also available from the NASA UHSAS, first corrected for the characteristics of BBA as described in Howell et al., (2021), for diameters up to 0.5 μm (the stated upper size limit for the correction algorithm. Concentrations are larger than those reported by any of the PCASPs. By converting the FAAM PCASP2 bin boundaries to those for BBA equivalent refractive index it can be seen that the PSD more closely matches that from the UHSAS although concentrations are still lower. This demonstrates the importance of considering the material refractive index when combining measurements from multiple probes with differing techniques.

14) Line 641. What are linear regression "sensitivities"?

Replaced "sensitivities" with "slopes" globally.

15) Line 664. Data taken at or below detection limit are also of use, and should be plotted as suggested in comment 1b above.

Data from the FAAM AMS are not available for the altitudes above the boundary layer during this flight. The instrument was not able to detect material above the background, and so can not be included here. A fit to these low magnitudes would be biased by data which is known to be of poor quality.

16) Line 688. "Re" (effective radius) is not defined.

Equation for Re (effective radius) is defined on line 487.

17) Line 677 (and 955). Show the LACIS ACMS data in Table 3. Are they at least correlated?

Added LASIC ACMS data to table 4 showing that LASIC ACMS always overreads compared to FAAM AMS, but by varying ratios.

18) Line 1080. Replace hyphen with a comma.

Replaced hyphen with comma.

References:

Please ensure that all references comply with Copernicus' style guide. For example, for Baumgardner et al. the title is capitalized, as is Cotterell et al. (2021). This behavior is a result of reference manager software, which always messes up formatting and must be thoroughly checked manually.

References checked and amended where required.

2. Response to Reviewer #2

We have included a key to acronyms as Table 8 and abbreviations have been checked.

2.1 Major

In general, comparing measurements with different setups, actively dried or not, is not recommended. To ensure comparable conditions, one should care for RH below 40 %. Especially the RH is of crucial importance for filter-based absorption photometers. The observed gradient in the RH (Fig 4c) transposes into the airplane's piping and will bias the absorption measurements due to the principle of differential measurement of the light attenuation behind the filter spots even if the cabin is heated to 30 °C (which also has implications for the volatile components of the aerosol particles). I.e., a sample at ~80 % RH at ~12 °C outside equals inside at ~26 % at 30 °C. As shown in the profile, there was a change to ~1 % RH at ~20 °C outside, which equals 0.6 % at 30 °C inside. This relatively fast change of more than 25 % can significantly impact the filter-based

absorption at NASA P3's PSAP or the TAP used on FAAM. However, the Nafion™ dryer at FAAM aircraft should dampen this effect significantly. The discussion must address this

feature of the experimental setup.

Relative Humidity is not controlled on all platforms: we agree that this is a significant issue – but in many ways it is this aspect that has motivated this study. The platform operators here (and in general) are very distinct, some operate state-of-the-art unique instrumentation - e.g. FAAM and EXSCALABAR for optical extinction, versus commercial instrument on NASA and LASIC, nephelometers and PSAP for optical scattering and extinction. We want to understand the comparability of measurements made using these techniques – in part to understand the comparability of our measurements across the SE Atlantic basin between 2016 and 2018 and also because a number of historical datasets already exist using a range of these techniques.

We include the profile plot of optical absorption and comment on the suspect artefact in

the PSAP sample from the elevated pollution layer. Added this to line 821:

"The FAAM PAS data from the profile descent shows that absorbing aerosols are present in magnitudes greater than the lower threshold of the instrument in the boundary layer, runBL, and upper pollution layer, runELEV. Data follow similar trends from the NASA PSAP in the boundary layer. In the elevated pollution layer the NASA PAS data look suspect, for example signals from red and blue are nearly identical suggesting an unphysical Absorption Ångström exponent (\AA_{AP}). This is likely because the PSAP is not suitable for operating in regions where pressure or RH or other external factors are changing rapidly such as during descent, especially, as is the case, where the sample is not actively dried. These data should be treated with caution and are not used in subsequent correlations (Fig. 6 (h), (i)). Consequently, the data for σ_{EP} , from NASA (nephelometer + PSAP) should be treated with caution in the elevated pollution layer, when compared against the FAAM CRDS measurement which probes optical extinction directly. "

We also have some discussion regarding RH already in section 5.4 which relates to the fact that the bias between LASIC and FAAM on the optical scattering measurements is in the opposite direction than might be expected from the un-dried LASIC sample. This continues into discussion around inlet sampling artefacts in section 5.5. We feel it is important to show these biases and consider the causes such that future campaigns may be better designed.

Table 3 is way too large. One should consider presenting the content more comprehensibly, like with figures. E.g., the table content can be separated into the coefficients of the linear fitting and average values.

We agree and have removed much of the material to the supplement, partly by including new Fig. 5

which compares temperatures and humidities and only keeping data which is not present graphically. We split the remainder in to multiple smaller and more targeted tables.

Figure 5 displays correlations of two variables consisting of uncertainty each. Hence a linear fit is not applicable, and an orthogonal fit accounting for both uncertainties should be applied. Moreover, it is unsuitable for fitting a linear behavior based on two observations. I would suspect that the statistical significance of those fits is small. Enhance the number of data points by decreasing the averaging window or address this in a deeper discussion.

We were originally using Orthogonal Distance regression fits to account for uncertainty/variability in both x and y directions, and this is now made clear in the text at the start of Sect. 4 Results*. We also take onboard the suggestion to reduce the averaging time (to 10s) where appropriate. This is done for the airborne comparisons. The fact that the data from the ground – airborne comparison are not collocated in space / time mean that this is not possible for this part of the comparison.

*Added line 615-634 - "When comparing measurements from two instruments, it is useful to explicitly consider statistical uncertainties, which differ between individual data

points, and systematic uncertainties, which affect all data points from an instrument. Statistical uncertainties are large when instrument noise is large compared to the measured signal, and/or the measured property exhibits a high degree of variability within the sampling period. The effect of instrument noise can be minimised by choosing a longer averaging time and this is the approach we take for the comparisons between the BAe-146 and ARM site. The straight and level runs were designed to minimise the variability of measured properties during the comparisons, and we average the data to one point per run. Conversely, where a large statistical uncertainty is caused by real variation in the measured property within the measurement period, a shorter averaging time must be used. This is the approach we use when comparing the BAe-146 and P3 aircraft, and here we average the data to 0.1 Hz to balance real variation with instrument noise.

Once a set of points for comparison has been gathered, we compare the variables using orthogonal distance regression (ODR) with results summarised in Table 3 and shown in more detail in the Supplement (sect. S7). These straight-line fits utilise the uncertainty in both the x and y variables (taken to be the standard error, equal to the standard deviation divided by the square root of the number of data points), to produce a fit uncertainty that accounts for the measurement uncertainty of each data point used to produce the fit. Comparison between the different platforms can then take place by comparing the slopes of the fits. Where they are different from unity both the statistical uncertainty of the fit and the systematic uncertainty in both instruments may contribute. When quoted in literature, this systematic uncertainty tends to be the calibration uncertainty, although other factors such as different inlets tend to make this uncertainty larger. Summary values of ODR fits for all parameters are to be found in Table 3. More completed tabulated results available in the Supplement (Table S2)."

Since a major point of the motivation is biomass burning aerosol, the discussion, and presentation of the aerosol particle light absorption coefficient is, in my opinion, not sufficiently addressed. Please also provide profiles of aerosol particle light scattering and absorption and a discussion of those.

We now include a plot of the profile of aerosol optical absorption, Fig 4 (g) for completeness. We also now include the profile of aerosol optical scattering. We suspect the NASA PSAP data to have an artefact of sampling, and we discuss this in the text. It is likely related to the nature of changing pressure on the sample flow to the filter and is the reason that we do not include data from the profiles in the subsequent analysis of aerosol optical absorption. Ideally we would have loitered at a fixed altitude once we located the elevated pollution plume, but it is not possible to do such changes to the planned flight path when flying in formation as we were.

2.2 Minor

Abstract

Line 40: please add ° in the coordinates

Added degree symbol

Line 52: first appearance: Avoid using "well" when comparing devices. Please rephrase.

Amended the text to remove ambiguous statements of "well".

Instruments

Line 115: Although referenced, no details on the SMPS of the AMS rack are presented in section 2.4.2

Added reference to SMPS in Sect. 2.6.

Line 121: exemplarily for other referencing parts in the manuscript. For all references, a period should adjoin the subsection. Instead of 2.52, it should read 2.5.2.

Changed the section numbering to have consistent format.

Line 118: Provide details of the CPC by referring to section 2.6, i.e., their volume flow rate.

Added CPC flow rates to Sect. 2.6, line 385-387, and reference to Sect 2.6 on line 119

Line 120 and 124: What means good? Within which range?

Added 10 to 30 % from the reference - "The inlet has been shown to efficiently transmit particles at dry diameters up to 4.0 μm (McNaughton et al., 2007) with good agreement for submicron sized scattering aerosols between this and ground based tower observations to between 10 and 30%."

Line 126: Please provide the particle losses due to the tubing as a function of particle diameter used to correct those losses, e.g., in the supplementary material.

These data were not corrected for modelled sampling losses due to pipework. Instead the sampling system was designed to minimise losses to a negligible level at the design stage of the instrument rack. The figures are not available here.

Line 132: Please provide the period of the periodical change.

Added details of PM10/PM1 switching regime.

Line 160: split ms⁻¹; otherwise, it is inverse milliseconds.

Corrected the units.

Line 186: Period after Beer's Law.

Added a period.

Line 209 and repeatedly appearing along with the text: Please avoid judgmental adjectives such as "good."

Many instance of this have been amended with reference to errors, uncertainties, or else rephrased.

Line 217: Remove one period.

Removed a period

Line 259: (first appearance): Ensure the optical coefficients are properly subscripted.

Corrected optical coefficients.

Line 300 and 359: Use a uniform notation; Nafion™ or Nafion(TM)

Corrected to a standard notation.

Line 354 and 359: Explain where the dilution of the aerosol arises and the underlying reasons. Comment in which why this was accounted for. Leakage of the Nafion™ membrane will bias the outside measurement with airplane cabin aerosol.

There is no accidental leakage, and aerosols can't pass across the Nafion(TM) membrane. Merely the instrument rack is designed such that the sample from outside is mixed with a clean filtered airstream, for reasons such as to provide a faster flow rate through instrumentation.

Line 378: (first appearance): AAE (absorption angstrom exponent) is not σ_{ap} . Please change.

Checked and corrected instances of absorption Ångstrom exponent.

Line 393: Comment or discuss where the factor of 1.8 originates from; Line 719: Comment on the underlying reasons for the empirical scaling factor used for the PSD.

Added commentary that details the processes of validating AMS volume concentrations with outboard PCASP, and then empirically scaling SMPS to better match the PCASP size distributions in the overlap region.

Line 400: According to the reference list, "Howell et al. (2020)" was published in 2021.

Corrected reference for Howell to (2021).

Line 428: Comment on the expected uncertainty omitting the refractive index correction of particles larger than 800 nm.

Added commentary to results section line 750: "A coarse aerosol mode was also present during runBL. At diameters larger than 0.5 μm , where particle counts are much lower, Poisson counting uncertainties become significant: 40 % at 1.5 μm and more than 200 % at 3.0 μm . The bin boundaries of the PCASP and CDP have not been corrected for the material refractive index, which is not known. The 2DS is a shadow imaging probe and so not affected by the refractive index of the material. Detailed scientific analysis should account for the materials refractive index and not doing so here does limit the utility of the results in the probe cross-over regions. However, the magnitude of the differences between PCASPs is much larger than the combined uncertainties at supermicron diameters. The largest differences are apparent between the two probes on the FAAM BAe-146 platform while FAAM PCASP2 and the NASA PCASP are in closer agreement. Only the FAAM CDP reported aerosol data in the particle diameter range 1-5 μm , but, at larger diameters, data from 2DS probes on both aircraft cross over with CDP observations and

show distributions with similar shapes. The cross over between CDP and PCASP is likely dominated by uncertainty in the larger sizes of the PCASP. This coarse mode will contribute to the total optical scattering from aerosol particles, as evidenced by the NASA runBL nephelometer data (Sect. 4.3.3) when switching between PM1 and PM10.”

Results:

Line 641: rephrase sensitivity to "the slope". Consistency: BAe-146 or BAe146. Choose.

Changed sensitivity to slope, and corrected to BAe-146.

Line 661: Discuss the differences in the measured CN between the two airplanes based on the cut-off of the CPCs.

Line 709 - Added discussion on lower cut-off diameter of CPCs

Line 792: One could update Figure 9, including the separation between NIR and VIS, and add the corresponding integrated values.

We feel that the diagram is suitable and note that the integrated values are presented in Table 8.

Line 900: Please comment on the volatile nature of ammonium nitrate evaporating

already at 20°C and its impact on the chemical composition measurements. See Schaap et al. (2004). Schaap, M., Spindler, G., Schulz, M., Acker, K., Maenhaut, W., Berner, A., Wieprecht, W., Streit, N., Muller, K., Brüggemann, E., Chi, X., Putaud, J. P., Hitzinger, R., Puxbaum, H., Baltensperger, U., and ten Brink, H.: Artefacts in the sampling of nitrate studied in the "INTERCOMP" campaigns of EUROTRAC-AEROSOL, Atmos. Environ., 38, 6487-6496, 10.1016/j.atmosenv.2004.08.026, 2004.

Added Line 980: “Ammonium nitrate is semi-volatile at atmospheric conditions and to investigate this a model of evaporation of aerosols to the gas phase was developed after Dassios and Pandis (1999) was run for a range of atmospheric conditions and a sample temperature of 30° C and a sample residence time of 2 s. This showed that the worst case scenario losses of aerosol mass to the gas was 7 %, assuming unity accommodation coefficient, instantaneous heating upon sample collection and a single aerosol component. Pressure and relative humidity exerted much weaker controls (< 2 %). Sample residence times may well be longer on the aircraft, but the uncertainty is related to the differences between the sampling set-ups on the aircraft rather than absolute values which also reduces the impact of this on the comparisons”

Line 1071: Provide a valuable reference for BBA density.

Line 1124 - Added reference to Levin (2010) for BBA density.

References

Add doi if available to each reference.

DOI added where available.

2.3 General Comments

Regarding tables: Table description on top of the tables.

The manuscript is very long. I recommend a revision in places that can be shortened. For instance, the instrument description part contains repetitive passages (e.g., gaseous components) and can be shortened, e.g., in the form of tables. A tabular overview of the instruments and corresponding parameters would be more understandable. After, differences between the airplanes and ARM-site regarding drying and instrument location (if necessary) can be explained.

Descriptions moved to top of tables.

Updating the colors of the fitting functions and adding the wavelength when optical coefficients are considered can improve figure 5.

Some of the parameters in the very long Table 3 are now plotted, allowing us to move those segments of the results table to supplementary materials. We did consider rationalising some more of the text in section 2 relating to instrumentation descriptions. We considered using a table to outline the instrumentation, then referring to that table in the text. However, although long, we feel that the section is well structured which aids understanding and readability and that the many bespoke details of the individual set-ups mean that much of the text would have to remain anyway. We felt that a slight shortened but still long text, allied to a table that needed referencing would not in the end assist the reader.

Figure 5, 6: Please provide the aerosol particles' volume and surface size distribution and their integrated and cumulative (along the diameter) sum values, e.g., in the supplementary material. Those would help comprehend the contribution of the different aerosol populations to the optical properties since those are a function of the cross-section of the aerosol particles.

Fig 5, (new Fig 6) this, and other figures have now been amended so that colours are used to distinguish that altitude of the measurements in most cases, or a particular instrument in others. We feel this has improved the figures. We have added the wavelength information where applicable.

Fig 5, 6 (new Fig. 6, 7), we agree that some further information on the particle size distributions was required. In conjunction with this comments and comments in Review 1 we opted to show the particle number and particle volume distributions from the airborne comparisons these show a wide range of conditions. This is added to new Fig. 6. Volume (and mass) are parameters that models such as general circulation models tend to represent as prognostic variables. Showing these parameters gives an overview of how particles across the size range are sampled in comparison to one another.

The area distributions are included in the supplement. The optical properties are hugely important and a large focus of this study. There is significant complexity in the optical properties as a function of particle size, e.g. most biomass burning aerosol is sub-micron,

and the composition of larger super-micron particles was not sampled. The optical properties depend strongly on composition and individual studies looking into these aspects of the science could be done, such as the study by Peers, et al. (2019).

We do not present cumulative distributions because we are relying on multiple probes to sample the full size range of aerosol particles. There is no obvious way to deal with the cross overs between individual probes and detailed study that produces a composite weighted fit is beyond the scope of the study. Likewise choosing an arbitrary size threshold at which to splice individual probes together would not be particularly instructive. Now we present both number and volume it is easier to see important features of the underlying aerosol size distributions.

Comment on the different observed size ranges of the different AMS systems, i.e., the difference between ACSM and AMS when comparing the chemical composition. I am not an expert in that field, but could it be that this explains the observed difference?

AMS and ACSM differences: There may be small size selection difference between the two instruments and sample inlets, of order 100 nm, but it is not envisaged that this is the driver of the differences. This is one set of comparisons that have been shown to be poor from this work, and unfortunately in this case we have not been able identify the underlying reasons.

Added line 1016 "The slight difference in quoted upper cut diameters of 600 nm (FAAM) and 700 nm (LASIC) do not explain these differences."

Line 1595: The specific instrument should be mentioned in the legend for each variable in all the figures. Change typo: its AAE (absorption angstrom exponent, not extinction angstrom exponent)

Individual instruments are now on the legend and typo has been corrected.

Figure 10a): Comment and discuss the discrepancy of one order of magnitude in the observed PSD of the 2DS and CDP.

Overlap between 2DS and CDP poor at small end of 2DS – added comment on large sample volume uncertainties.

3. References:

Levin, E. J. T, McMeeking, G. R., Carrico, C. M., Mack, L. E., Kreidenweis, S. M., Wold, C. E., Moosmüller, H., Arnott, W. P., Hao, W. M., Collet Jr., J. L and Malm, W. C.: Biomass burning smoke aerosol properties measured during Fire Laboratory at Missoula Experiments (FLAME), *J. Geophys. Res.*, 115, D18210, <https://doi.org/doi:10.1029/2009JD013601>, 2010.

Peers, F., Francis, P., Fox, C., Abel, S. J., Szpek, K., Cotterell, M. I., Davies, N. W., Langridge, J. M., Meyer, K. G., Platnick, S. E., and Haywood, J. M.: Observation of absorbing aerosols above clouds over the south-east Atlantic Ocean from the geostationary satellite SEVIRI – Part 1: Method description and sensitivity, *Atmos. Chem. Phys.*, 19, 9595–9611, <https://doi.org/10.5194/acp-19-9595-2019> , 2019.