

Response to Reviewers

K. Kandler (Referee #1)

The manuscript describes airborne measurements of mineral dust size distributions close to the source. Representation of aerosol size distributions inside atmospheric transport models is still based on many assumptions, and new data for evaluations and input is highly welcome. The manuscript adds new data for the region close to the source at the beginning of the trans-oceanic transport. It also adds additional evidence to the importance of studying particles larger than 10 μm , which are frequently omitted in aerosol research, and which are difficult to study. Furthermore, optical properties for the dust aerosol are reported, which can also serve the community, e.g., as model input or for remote sensing interpretation. The authors show that variation in physical as well as compositional parameters can have a considerable impact on the optical properties, and that in their samples, the compositional variation is dominating the variation in single scattering albedo.

The topic is suitable for ACP. The paper is well-written and clearly structured. References are mostly made where appropriate, some additional comparison with previous work would enhance the general significance. Some issues - particular on the error handling - should be addressed and questions answered before publication.

We would like to thank Konrad Kandler for this detailed review and comments which we hope have led to an improved, clearer paper. We have added information regarding the different sizing metrics used, additional explanations of error handling, and expanded on comparisons to previous work, all of which are detailed below. We apologize for incorrect cross-referencing of tables and figures, which have all now been corrected.

=====

Remarks and questions

P3 L32: Are really any inlet size restriction removed? Given the high speeds at which particles are collected, a minimal deviation from perfect alignment might easily introduce boundary layers and re-circulations.

It is true that variations in the airspeed and flow angle through the sample area may occur due to distortions of the airflow about the wing probe housing and arms (Korolev et al. (2013), Weigel et al. (2016), McFarquhar et al. (2017)). We do not attempt to correct for these factors, and now state so in Section 2.3.2, and mention them in the introduction.

In terms of the alignment, the canisters are set at a set vertical angle to accommodate the 'normal' angle of attack (AoA) of the aircraft. There will be up to several degrees of misalignment as the AoA changes with altitude. The average and standard deviation of the canister angles is 3.25 ± 0.14 deg (negative is nose-down). The range of AoA is 3.5-6 degrees nose up so there will generally be a couple of degrees misalignment between the probes and the airflow assuming no aircraft perturbation of the flows. Korolev et al. (2013) show that the velocity between the arms of the CIP has an unperturbed region of approximately half the arm width. This suggests that a small misalignment (only several degrees) is unlikely to narrow this unperturbed region sufficiently to impact the size-dependent concentrations.

Measurements behind inlets will systematically omit a portion at the upper size range of the PSD, and while wing probes may be subject to various uncertainties relating to local flow distortion, they are at least able to measure the full size range of dust particles because no inlet is used. Additionally, the bigger the particle the less they are effected by flow distortion because of their larger momentum. This is the opposite of the pipe bend problem which imposes a cut size.

P6 L32: Should this diameter then considered as optical equivalent diameter? Please state.

Not in this case, since various ‘types’ of diameter are considered throughout the paper. However, we have added, “Hence particle sizes determined from the PCASP and CDP represent optically equivalent diameters,” to section 2.3.2 where this is discussed.

P7 L10: Have experimental or theoretical approaches been used to characterize the losses?

Both – Trembath (2012) performed OPC measurements behind different inlets on the aircraft, and followed these up by theoretical calculations of pipe losses. We added “experimental and theoretical” to this sentence.

P8 L6 and Figure 2: It looks like the sizing uncertainties from the number size distribution were not propagated into the volume size distribution (e.g., the CDP point at 20 μ m seems to have a factor of 3 uncertainty in size, which should propagate to around a factor of 10 in volume, but a much smaller error bar is shown in the volume plot). Instead it looks like the same vertical error bars are shown in both plots. Please explain or modify.

The error processing takes place in two stages, which may be easier to follow separately. 1) Uncertainties are calculated for each SLR. 2) Uncertainties for the campaign average PSD are calculated, which are shown in Figure 2. We explain each of these steps below.

Additionally, the horizontal error bars on Figure 5 represent the minimum and maximum bin edges possible, using the bin centre and bin width uncertainties. Hence the actual uncertainty on the diameter centre point (which would be used in a typical progression of uncertainties from number concentration to volume concentration as stated by the reviewer), is much smaller than that shown on the plot. We show error bars to incorporate both bin centre and bin width in order to provide a realistic reflection of the real uncertainty in the sizing.

1) For each SLR, $dN/d\log D$ uncertainties (vertical error bars) already includes bin width uncertainty *and* bin centre uncertainty, since they both contribute to the calculation of $d\log D$. $dV/d\log D$ additionally takes account of the bin centre uncertainty again through the diameter cubed, as well as incorporating the uncertainty in $dN/d\log D$. Below is an example from one SLR, b932 R2, in the SAL.

In the figure below, we take the reviewer’s data point close to 20 μ m (diameter 18.7 μ m). Here the uncertainties are as follows:

Bin centre uncertainty = 0.74 μ m

Bin width uncertainty = 0.71 μ m

Fractional uncertainty in $dN/d\log D=0.28$

For the uncertainty in $dN/d\log D$, the uncertainty from the bin width is 4 times larger than that from the uncertainty in the measured number concentration, so that the bin width error dominates the total error.

Fractional uncertainty in $dV/d\log D=0.29$

For the uncertainty in $dV/d\log D$, the uncertainty from the bin centre is around half that of the uncertainty propagated from $dN/d\log D$ (which also already incorporates both bin width uncertainty and bin centre uncertainty). As a result, the fractional uncertainties in $dN/d\log D$ and $dV/d\log D$ are not very different in the two plots below, since they are both dominated by uncertainties in the bin size.

In contrast, for the CDP data point centred at 3.9 μ m, the fractional uncertainties in bin centre and bin width are much larger (0.12, 0.05), which results in the much larger vertical error bar in both $dN/d\log D$ and $dV/d\log D$.

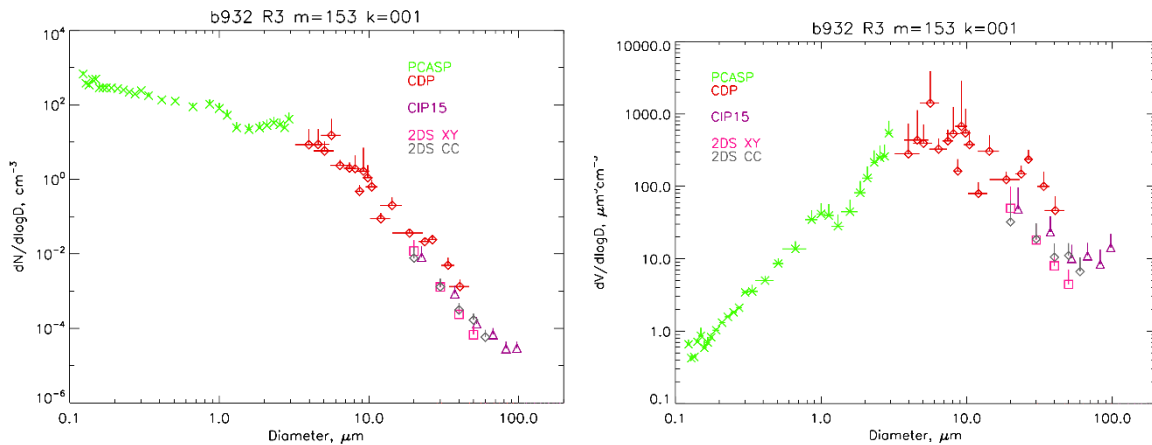


Figure: $dN/d\log D$ (left) and $dV/d\log D$ (right) for an individual SLR (b932 R3) in the SAL.

2) Secondly, uncertainties are combined to provide the SAL and MBL average PSD, e.g. as shown in Figure 2.

In order to do this, the random and systematic errors from each SLR are processed separately, as recommended by Baumgardner et al. (2017). This is because the systematic error (from sample area uncertainty and bin size uncertainty) does not reduce by increasing the sample size, while the random errors (from counting and discretization uncertainties) can be reduced by increasing the sample size. This is done by standard analytical error propagation. It turns out that for the CDP, beneath 20 μm diameter, the total uncertainties are dominated by the bin size uncertainties, while above this size the counting error becomes significant as well.

We have added information to the caption for Figure 2, “Horizontal error bars represent maximum bin width due to uncertainties in both bin centre and bin width.” We have also added the following text to Section 2.3.2:

“For the CDP at $d < 20 \mu\text{m}$ bin size uncertainty was found to dominate the total uncertainty in $dN/d\log D$ and $dV/d\log D$. Horizontal error bars in Figure 2 represent the maximum uncertainty in bin edges, derived from uncertainties in both bin centre and bin width. Uncertainties in bin size contribute to uncertainties in both $dN/d\log D$ and $dV/d\log D$, and therefore the relative uncertainties do not change significantly between the two panels.”

Figure 2 is hard to understand also from another aspect: Either the measurement points are all at the lower boundary of the error, or the error always includes zero. If they are at the lower boundary, please explain why. In particular with a non-zero counting error (which is necessarily two-sided), this seems impossible. If the error always includes zero, then the statement in P9 L27, “agree within the error bars” is meaningless, as all data could be zero.

In Figure 2 (and also Figure 5), only upper error bounds are shown, since for several points the lower error bars exceed the minimum on the log scale axis and impede visibility of the data points. This is now stated in the caption for both figures.

P8 L31 and P9 L4: if a particle was not counted for the given reasons – how was this underestimation of the concentration then accounted for? Decreasing proportionally the sample volume?

Yes, it is accounted for. The effective array width (EAW) changes depending on whether the all-in or centre-in approach is used. The EAW is used in calculation of the sample area (McFarquhar et al., 2017), and therefore the number concentration and its uncertainties. We added, “The sample area is adjusted for the effective array width, which is different depending on whether ‘all-in’ or ‘centre-in’ is used (McFarquhar et al., 2017), and therefore the calculated number concentrations account for this.”

P9 L6: For what reason the projected area equivalent diameter was not used, which appears to be a common standard in shadowing techniques (i.e. light microscopy)?

With hindsight, and ample data processing time and computation time, it would have been interesting to additionally process the OAP data using projected area equivalent diameter. However, a limitation of using this metric with the OAPs is that particles can sometimes appear hollow (e.g. due to being out of focus). Under these circumstances using projected area equivalent diameter would be a significant underestimate, and therefore it is not clear that this is the 'best' metric for the OAPs. We have added the words, "Area-equivalent diameters were not calculated because particles can sometimes appear hollow on the OAPs (McFarquhar et al., 2017) which would lead to undersizing."

P9 L11-21: Please be more explicit on the error handling. Was it done by analytical error propagation, by Monte Carlo simulation, ...? What were the distribution assumptions (Poisson?)? How was the counting error treated, if measurement points with only 4 particles were regarded? If error bars are given: do they show a certain confidence interval, one (or more) standard deviations, or maximum errors?

Error handling was done by standard analytical error propagation. Counting error was calculated as $n^{1/2}$, where n is number of particles counted in the SLR. Size bins containing less than 4 particles in the whole SLR were disregarded, and therefore counting error was not calculated for these points. Error bars show maximum errors. This paragraph now reads:

"Uncertainties in number concentration for all size probes are propagated from 1Hz measurements, through to means over SLRs, through to the AER-D campaign averages. For all probes, random errors (due to counting and discretization error) and systematic errors (due to sample area uncertainty and bin size centre and width from Mie singularities) were accounted for in their contribution to total number concentration errors, and propagated by standard analytical error propagation. I.e. random error can be minimized by increasing the sample size (averaging across the campaign), while systematic error remains constant. For the CDP at $d < 20 \mu\text{m}$ bin size uncertainty was found to dominate the total uncertainty in $dN/d\log D$ and $dV/d\log D$. Horizontal error bars in Figure 2 represent the maximum uncertainty in bin edges, derived from uncertainties in both bin centre and bin width. Uncertainties in bin size contribute to uncertainties in both $dN/d\log D$ and $dV/d\log D$, and therefore the relative uncertainties do not change significantly between the two panels. All error bars represent maximum uncertainty."

P9 L30-35: Are there any arguments beyond consistency with previous measurements for disposing (if I understood correctly) the CIP15 CC data? It is stated that the 2DS XY metric is better – better in terms of what? Please explain.

Yes, as stated on p9 L29, the "the CC metric will oversize a non-spherical particle," since the circumscribing circle, by definition, does not take any account of a particles' non-sphericity. We refer specifically to the metric, and not the instrument here.

P10 L1: This type of effective diameter seems usually (Hinds 1999) to be called the Sauter mean diameter (apparently defined by Sauter 1926) or mean volume-surface diameter (Hinds 1999). As different effective diameters exist, I suggest referring to one of the mentioned denominations and referring in addition to a more fundamental paper (McFarquhar et al. 1998).

We have changed the effective diameter reference to Hansen and Travis (1974), and added the Hinds (1999) reference as well as adding 'volume-surface diameter.'

P10 L4-7: Given the high counting uncertainty associated with a single count of the maximum-sized particle, wouldn't there be a better option, e.g., fitting a log-normal distribution or power law into the data? At least, both mentioned approaches should be compared to allow for consistent comparison.

Since we disregard SLRs where fewer than 4 particles were detected, our counting uncertainty will be a maximum of 50% if 4 particles are detected, and reduce accordingly when more are detected. Actually, only 2 SLRs detected fewer than 6 particles in the maximum size, so the errors are substantially less. Fitting a lognormal or a power law to the data is not necessarily useful as the size distribution shape will not necessarily follow these fits.

P11 L12 (and below): As magnifications are relative to the used screen, a measure like nm per pixel or pixels per smallest particle would be more meaningful.

Values are: x2,000: 55.9 nm/pixel, x10,000: 11.0 nm/pixel. These have been added to the manuscript.

P11 L14-19: Wouldn't it make more sense using the same metrics as described for the shadowing instruments above?

In an ideal world, yes. The OAP software processing is computationally intensive, and ellipse fitting to the time-of-flight data is not currently an option. Since more realistic metrics could be derived from the filter samples, we chose to use these (rather than a less appropriate metric consistent with the OAP data), hoping to inform knowledge on dust properties for the dust aerosol community. In the future, with the benefit of hindsight, processing the OAP data differently would be considered. See also response to P9 L6 above.

We tested the impact of processing the filters PSDs differently, mimicking the XY and CC metrics as much as possible. However it should be noted that even when processing the filters data like this, the two are still not equivalent. This is because the 2DS and CIP15 data are 2-D projections of a 3-D particle orientations in the atmosphere. While the filters data are also 2-D projections, the particles likely fall with their largest axis parallel to the filter sample. Despite this, using the XY metric for the filters data did not produce any noticeable changes in the PSD. Contrastingly, the CC metric processing produced PSDs with a mode shifted towards larger particles, and often detected a maximum particle size one size bin greater than that shown in Figure 7. These effects on the PSD mimic the differences seen between the 2DS CC and 2DS XY in Figure 2.

We have added a sentence to Section 2.4 to indicate this:

"Additionally, we tested the sensitivity of the filters PSD to using the mean XY and CC sizing methods applied to the OAP data (not shown). Using a mean XY method on the filters data did not produce significantly different results, while using the CC method was found to shift the PSD towards larger particles, similar to the findings from the OAP size metric comparisons."

P11 L25-29: Please state where you draw the lines between thenardite/sulfate and gypsum/sulfate.

No quantitative separation is used since the percentages vary depending on the contribution from the other oxides. However, typically when Thenardite (Gypsum) is classified, percentages of Na₂O and SO₃ (CaO and SO₃) are within 10% of each other. We have added information about the classification of Thenardite to section 2.4 which was missing, and added information on the oxide percent.

P16 L18-20: This behavior was described and a similar explanation was given for the same geographical region and season a while ago (Jaenicke et al. 1978), and also some model approaches seem to predict that (Garrett et al. 2003), so maybe a short comparison with previous finding would make sense here and enhance the general significance of the findings.

Thank you for these interesting references. We have added to this section the following, "This has also been suggested by Jaenicke and Schutz (1978) from aerosol surface observations at Sal, Cape Verde, where giant particles ($d > 40 \mu\text{m}$) were observed to arrive at the site a day after that of coarse dust particles ($6 < d < 60 \mu\text{m}$)."

Garrett et al. (2003) appears to be more relevant to the introduction, since that work found that concentrations of $d > 10 \mu\text{m}$ particles were under predicted by models for transport across the Atlantic, so we have added that there.

P16 L28-30: This finding is surprisingly similar to the volume distributions shown for Cape Verde in winter time, comparing dusty and marine situations (Kandler et al. 2011), so a comparison could show a broader relevance.

Thank you for this reference. Actually this finding is a little different from those of Kandler et al. (2011). In Kandler et al. (2011), for the dust case mean, the volume distribution peaks sharply at around 10 microns, whereas for our 'giant dust in the MBL' cases the volume distribution peaks at $\sim 5 \mu\text{m}$ and $\sim 30 \mu\text{m}$. For the 'maritime' cases the PSDs shapes appear quite different too.

We have added a paragraph to Section 3.2 comparing the PSDs to those in the literature. We are also about to submit a follow-up paper which includes a thorough comparison of the AER-D PSDs to those in the literature, drawing together the latest results.

P17 L14-18: This would mean that the flow through the filter would have been ten times as high as the one used for calculating the size distributions – is this still in a physical probable range? Can there exist any aerosol concentration effects due to high velocity gradients during sampling?

Firstly, yes, this is plausible. There is nearly a factor of 10 difference in flow rate measured between different SLRs. Therefore it is possible that problems here could have strongly impacted the retrieved size distributions.

Secondly, non-isokinetic sampling can artificially increase number concentrations in certain size ranges. E.g. if the aspiration speed is larger than the aircraft speed then the submicron fraction will be artificially enhanced, and vice-versa. However, this is not consistent with the filters PSD being offset across all size ranges compared to the wing probes.

P17 L22: Smoothness of the curves could be related to different size intervals. The wing probes seem to have more size intervals (with higher counting uncertainties) in Figure 7.

This is true. We have added, “in part due to the broader size bins used” to this sentence.

P18 L23-32: Probably a median aspect ratio, instead of a modal value, would make interpretation less dependent on a single interval. In particular, as in the following the values are compared with median values from other sources, which otherwise can't be compared.

Thank you for this suggestion. We have calculated median values of the aspect ratios and added them to the panels in Figure 9. We have incorporated them into the discussion in this paragraph and re-written it. Although calculation of the median reveals higher values than the modal values, our values still appear a little lower than those from SAMUM1, 2 and AMMA, ranging from 1.16 to 1.54 for the full PSD.

P20 L24-25: How about sea-salt reacted with sulfuric acid?

Yes, this has been added.

P21 L7-12: How would black carbon have been identified in this work? Were all particle images manually inspected for fractal-like structures (doesn't come clear from the method section).

Yes, all images were manually inspected during the scanning procedure. Out of 6500 particles analysed, only one single black carbon particle was observed, in a chain, or fractal-like structure.

We have adjusted this text to reflect this more accurately, which now reads, “Additionally, in contrast to Liu et al. (2018), we do not detect any black carbon on the filter samples **in significant quantities**, which they find present predominantly between sizes of 0.1 to 0.6 μm . **During the analysis of 6500 particles, only one black carbon chain structure was observed.**” We also added the following sentence to the methodology section 2.4 for clarification, “Only one black carbon chain-like structure was observed during the analysis of over 6500 particles, and therefore this aerosol category is not included.”

P22 L20: In particular for the internal mixing, there is plenty of room with respect to complexity for calculating an effective value (Nousiainen 2009; Lindqvist et al. 2014).

These citations have been mentioned in this sentence now.

P24 L6: In addition, also measured refractive indices show the dependency on iron (Moosmüller et al. 2012; Caponi et al. 2017), so it appears to be consistent in general.

We have added these citations to this section.

Figure S2: Where does the ‘step function’ of the imaginary part at 700 nm wavelength derived from?

This 'step function' at 700 nm originates from the refractive index dataset of hematite, which drops sharply from ~ 0.2 to ~ 0.0013 in the imaginary part at this wavelength.

=====

Corrections

General: The order of the figures does not correspond with the order of their references in the text.

General: The table numbers in the text don't fit with the table numbers at the end.

Apologies for these errors, these have now all been corrected.

P3 L8-11: including into calculation/model?

This was calculated – the wording has been changed.

P4 L22: Why time-of-flight? Wasn't it an SP2 instrument?

We have changed the wording to 'real time measurements,' which is more accurate. (Yes, it was an SP2).

P4 23-25: While the link to the optical property measurement is clear, the reference to the ice nucleation measurements and to the modeling work doesn't seem to add something useful here.

We believe it is useful to tie together the papers relating to the same flights since the impacts of dust in the atmosphere are multiple. However, we have deleted the reference to the paper in preparation since it is not yet submitted.

P4 26-30: I suggest removing the paragraph, as the structure is standard.

Done

P5 L15-22: If results from SAVEX and CATS are not discussed here, these explanations should be removed.

These flights and results are discussed here. Certain flights covered multiple objectives. We prefer to retain this information since it explains the choice of location for each of the flights presented in this article and may be used in subsequent publications.

P8 L2: "ambiguities" instead of "singularities"?

Changed.

P9 L9: What does "Instead, ... sizing metric" refer to? Aren't the previous sentences the section about the sizing metric?

Here we intended to refer to the sizing metric of XY versus CC, pointing out that the choice of XY vs CC is the main controller of PSD in this case, as opposed to the choice of an 'all-in' versus a 'centre-in' approach. We have clarified this by adding, '(i.e. XY versus CC)' to this sentence.

P10 L8-9: "do not have a minimum detection concentration level, but at low particle concentrations the sampling statistics simply become poor": In fact, this is the detection limit. With poor sampling statistics the (counting) error becomes high in comparison with the signal. If it is decided to omit data below a certain signal/error ratio, the detection limit is introduced (which is a common procedure).

This is actually what we intended to say. To clarify, we have added to this sentence, ", introducing an effective detection limit."

P10 L14: "PSD' size distribution" – doubling "size distribution"

Removed

P12 L10: “low SEM signal” low image contrast?

By ‘low signal’ here we refer to the energy, or number of photons emitted, not the image contrast. Since this is dependent on particle volume, for smaller particles the number of photons emitted is lower, or non-existent for very small particles, and the analysis and interpretation becomes difficult. We have added, “(fewer photons emitted for smaller volume particles)” to this sentence.

P15 L5: increased?

Changed

P16 L24: There seems to be no (b) in Figure 6.

Changed

P18 L5: In Figure 7, in most cases the volume maximum is below 10 m, in one above (filter size distribution). “Dominating” therefore doesn’t seem to be appropriate for the largest particles with respect to the mass.

This has been changed to, “can contribute substantially to”

P18 L12: Sentence ends with a comma.

Changed.

P24 L28: Most of section 4 is rather a summary than conclusions, so the section should be termed accordingly.

We have followed the guidelines from ACP at https://www.atmospheric-chemistry-and-physics.net/for_authors/manuscript_preparation.html for manuscript composition, which indicates that ‘Conclusion’ should be used. Also in our final paragraph we make some broader recommendations and comments.

Table 1: Different reference formats. Maybe explain instrument abbreviations, e.g., as addition in Table 3?

The reference formats have been corrected. We have created an appendix containing the instrument acronyms, and added this to the caption of Table 1.

Table 3: Check caption

This has been corrected.

Table 4: different time formats between table 2 and 4 (?)

Table 4 has been changed to comply with the given ACP date and time terminology guidelines.

Table 5: ‘IT’ number format (E+01 etc.)

These have been changed to the 1.16×10^1 format.

Figure 2: Both y axes should have the same number format.

The axis for the right hand panel has been changed to be consistent.