On the issues of instrument performance and shattering artefacts for the FSSP and CIP

Since all four reviewers raised the issue of shattering (and instrument performance) as major criticism we decided to provide a separate reply addressing only this topic. At first we present our arguments and in the end we detail what changes have been implemented in the revised manuscript in order to highlight the issues and their possible influences on the presented data such that a reader can arrive at his/her own opinion.

1. General arguments:

1.1. Wind tunnel experiments by A. Korolev and Korolev et al., 2011, BAMS paper: The wind tunnel studies as well as the measurements underlying the 2011 BAMS paper were conducted under vastly different conditions when compared with the measurements we report on from the West African MCS. Our measurements were performed at much lower temperatures (i.e. below -40°C and down to -80°C) and at much lower humidities and number concentrations than (1.) the wind tunnel measurements and (2.) the measurements in mostly mixed phase clouds of the BAMS 2011 paper. Thus, an extrapolation of the factor 100 to 1000 enhancements due to shattering from the reported conditions of the BAMS paper to our low temperature, low humidity, low concentration conditions is not justified until solid evidence for this is provided for example by dedicated wind tunnel experiments or more airborne instrument intercomparisons (like e.g. Jensen et al., 2009) at such UT/LS conditions. This is also suggested by the study of Lawson (2011; Effects of ice particles shattering on optical cloud particle probes; AMTD, 4, 939-968), who performed measurements in anvil cirrus at temperatures from -30°C to -63°C and concluded that the post processing interarrival time analysis is well suited for rejection of suspected shattered particles and that this analysis might even carry more weight than the application of modified tips. For example, the ice particles we measure have no quasi liquid layers and may even consist of a glassy physical structure (Murray et al., 2010) and thus their break up and bounce behaviour will differ from the behaviour of the much warmer hydrometeors. For these reasons we think that it is by far premature to conclude that ***all*** FSSP data are obsolete because of the shattering influence, especially when other data from partly overlapping instruments -like the CIP- are available. We also believe that careful analyses of the individual size distributions still can provide useful data. What our “careful analysis” implies which ultimately led to the rejection of data not included in the manuscript, is described below in Section 2 and 3. Of course, we cannot exclude shattering influence completely, but limit the data of our publication to cases where we believe such influence is low, and highlight the possibly contaminated data to the reader.

1.2. CIP and FSSP performance: In the paper by Jensen et al. (page 5523, On the importance of small ice crystals in tropical anvil cirrus, ACP, 2009) it is stated: “The agreement between size distributions derived from the CIP and 2D-S imaging instruments is excellent throughout their overlapping size range (>50µm).” The authors refer to tropical measurements in Costa Rica at 11.4km to 12km altitude and probably these are similar to the “young outflow” conditions of 7 August 2006, in Section 4.2 of our manuscript.

Concerning the FSSP, Cairo et al. (A comparison of light backscattering and particle size distribution measurements in tropical cirrus clouds, AMT, 4, 557–570, 2011) state in their conclusions: “A comparison of optical properties for tropical high altitude cirrus clouds, directly measured and inferred from particle size distribution observations, has been carried out. Results suggest that the fraction of the size spectrum available from FSSP particle counter observation, i.e. particles with diameters from 2.7µm to 31µm, is effective in reproducing cirrus optical properties in the visible part of the spectrum. This result keeps validity for backscattering cross sections spanning over 5 orders of magnitude. Optical particle counters observations are thus a valid tool to assess the cloud particle density and to provide size distributions for modelling cloud microphysical processes and radiative effects in the visible region of the spectrum." To arrive at this result Cairo et al. used a careful selection of our SCOUT-O3 data from Darwin, Australia, our SCOUT-AMMA data from West Africa, and our TROCCINOX data from Brazil. They applied the MAS backscatter sonde on Geophysica to measure directly backscatter and depolarisation in the vicinity of the aircraft within the tropical high altitude cirrus clouds and compared this data with the backscatter ratios derived from the *in situ* measured particle size distributions from the FSSPs. The optical backscatter for the most part depends on the small particle concentrations reported by the FSSP and much less on the larger sizes from the CIP. If shattering had enhanced the corresponding small particle number densities by factors of 100 to 1000, then this intercomparison would have severely failed. Of course, this only holds for the used data sets and may not be "extrapolated" to cirrus in general.

1.3. Gas phase derived IWC vs. CIP&FSSP: In our manuscript we refer to the curve below as published in de Reus et al., ACP, 2009. Here, the IWC derived from Lyman-α hygrometer H_2O measurements (on the ordinate total water minus gas phase water) are directly compared with the concurrently measured particle IWC (abscissa), as calculated by using the Baker and Lawson (2006) scheme:

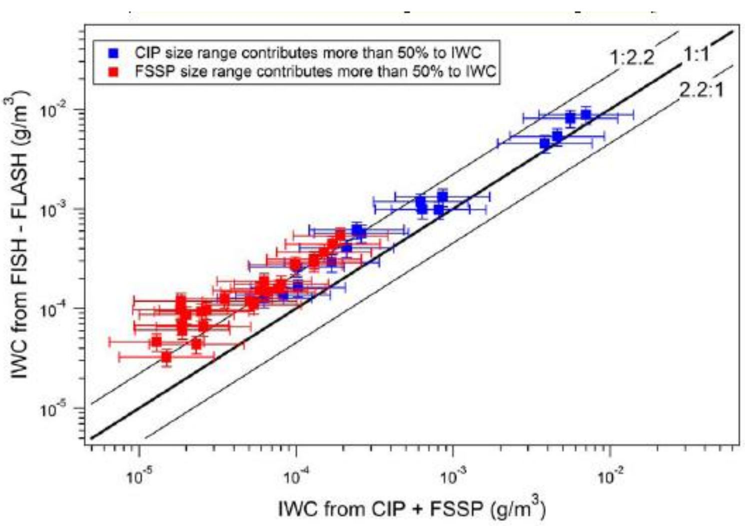
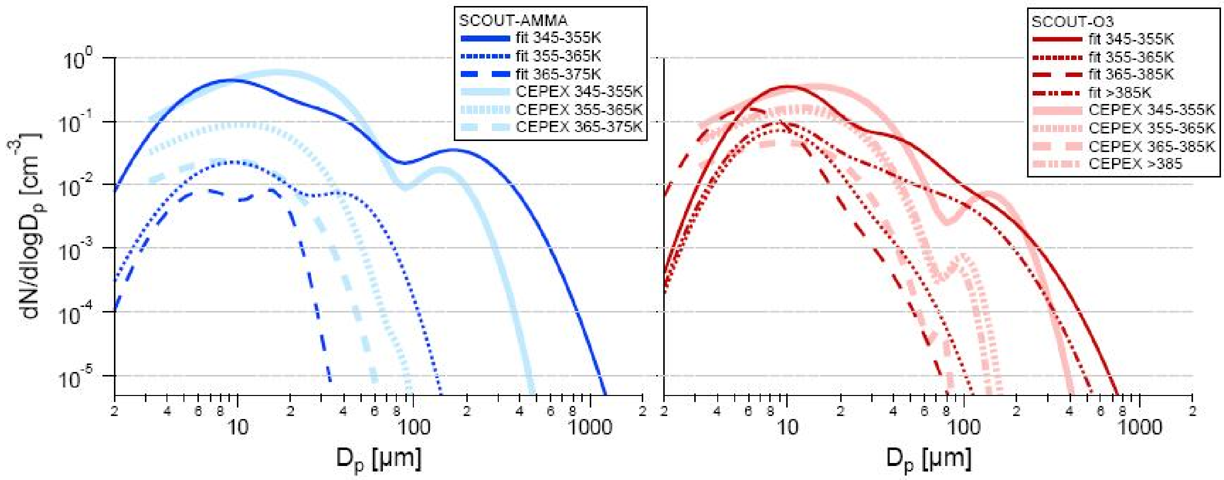


Fig. 4. Comparison of the ice water content derived from the size distribution measurements to the IWC obtained from the hygrometer instruments. The thin solid lines represent a deviation from the 1:1 line (thick solid line) with a factor 2.2, which corresponds to the combined uncertainty of both methods. The data points are averages over at least 30 s and correspond to the ice cloud encounters on 29 November and the two flights of 30 November from Fig. 3.

Anonymous Reviewer #4 remarks that the underlying volumes mostly are influenced by the large particles, which is true for IWC larger than roughly 0.001g/m³. For the smaller IWC the FSSP size range contributes more than 50% of the IWC. At least here one can assume that additionally detected particles from shattering would enhance the IWC artificially for the CIP&FSSP on the abscissa. However, in the graph the IWC from the particle instruments are even *too low* when compared with the gas phase instruments. The IWCs of the encountered outflow events as presented in our manuscript range from $6 \cdot 10^{-6}$g/m³ to $6 \cdot 10^{-2}$g/m³ with many values below 0.001g/m³. For the data presented from the subvisual cirrus and uppermost UT cirrus the IWC were much lower than 0.001g/m³. (Note for clarity: A similar plot unfortunately cannot be prepared for the AMMA flights, because the four involved instruments were not often enough operational concurrently at the same time and while inside the clouds.) At least for these cases it is unlikely that shattering influence of factors between 100 and 1000 would have gone by unnoticed.

1.4. Comparison AMMA/SCOUT data with CEPEX: In Figure 1 of the manuscript a summary of our SCOUT-O3 data from Darwin and the SCOUT-AMMA data from West Africa is presented together with the parameterisation which McFarquhar and Heymsfield extracted from their CEPEX measurements. In the lowest potential temperature bin of Figure 1 IWCs larger than 0.001g/m³ were found such that the argument from Section 1.3 is not applicable here. However, our measurements agree quite well with the CEPEX parameterisation *particularly* for the particle sizes below roughly 20µm. The major differences between CEPEX and SCOUT-AMMA occur at the very large sizes. During CEPEX the particles were measured with a VIPS and a 2-DC probe. The VIPS has an entirely different “inlet”-geometry and measurement principle w.r.t. the FSSP and shattering -presumably- is not an issue. If shattering had introduced artefacts to our particle number concentrations here on the factor 100 to 1000 levels like indicated by the BAMS 2011 paper, then this intercomparison would have turned out very differently.



This point is better visible in the figure above from the PhD thesis of Wiebke Frey, 2011. Again the log-normal fits of our measurements compare well with the CEPEX parameterisation, *especially* for the small sizes in the lowest potential temperature bin. (For the higher potential temperatures our results lie well below the CEPEX data and/or particles were too small for significant shattering.) Here, the ordinate displays absolute concentrations and not normalised as in Figure 1 of the manuscript.

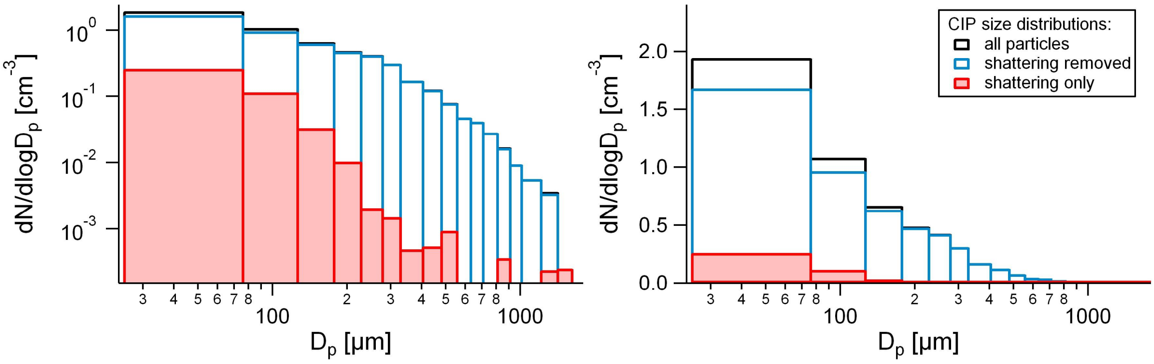
Based on these general arguments we believe that careful inspection and selection of the data does allow us to retain the CIP data and some of the FSSP measurements.

2. Shattering artefacts for the CIP and small particle detection:

2.1. Small particle detection: Reviewer #4 raises the point that there may be problems with the CIP for diameters less than 150µm. Our CIP has newer electronics with faster response times compared to the first CIP instruments and the 2DC probe. Baumgardner et al. (2001, The cloud, aerosol and precipitation spectrometer: a new instrument for cloud investigations. Atmos. Res., 59, 251-264) described this improvement and determined that there is no more dependency of the depth of field on the aircraft's velocity (which was described by Baumgardner and Korolev 1997, Airspeed corrections for optical array probe sample volumes. JAOT, 14, 1224-1229). Furthermore, Lawson (2011; Effects of ice particles shattering on optical cloud particle probes; AMTD, 4, 939-968) stated, based on recent measurements, that the CIP reliably measures droplets of 50µm at speeds below 150m/s. During our measurements the aircraft velocities were 135m/s (young outflow case), 140m/s (recent outflow case), and between 145m/s and 157m/s in the aged outflow. Thus, we assume that the CIP has reliably reported the particles over its entire detection range, keeping also in mind the statement from Jensen et al. (2009, see Section 1.2 above).

2.2. Interarrival time criterion for shattering of the CIP: We agree that, if the interarrival time threshold is set to too high values, a significant fraction of legitimate particles may be removed. If it is set to too low values, a significant part of shattered particles may not be removed. Therefore, the interarrival time threshold was determined specifically for each particular flight. In our analyses of the SCOUT-AMMA data the interarrival time rejection has always been applied to the whole data set of each flight including the subvisual cirrus cases. The shortest interarrival time encountered by the CIP measurements during the four SVC cases was 180µs which is much larger than the interarrival time thresholds of 2.6µs to 5µs which were adopted for the respective data sets. For this reason no particles were erroneously removed from the SVC cases and we can unfortunately not perform the test in the way suggested by reviewer Darrel Baumgardner. The shortest interarrival time encountered during the aged outflow events from 11 August was 1300µs and thus, no particles have been rejected during these events.

2.3. Percentage of shattering for the CIP data: The fraction of shattered particles is given as percentage of the total number of particles detected by the CIP. As reviewer Grant Allen states, this percentage does not give information about the particle sizes and correspondingly we provide here additional information. Over 85% of the shattered particles are found in the smallest three size bins (i.e. particles smaller 175µm). Shown below is the CIP size distribution of outflow event 1 (7 August 2006) from the manuscript, where 10% of the particles detected by the CIP were shattered fragments. The shattered fragment particles are displayed in red, while



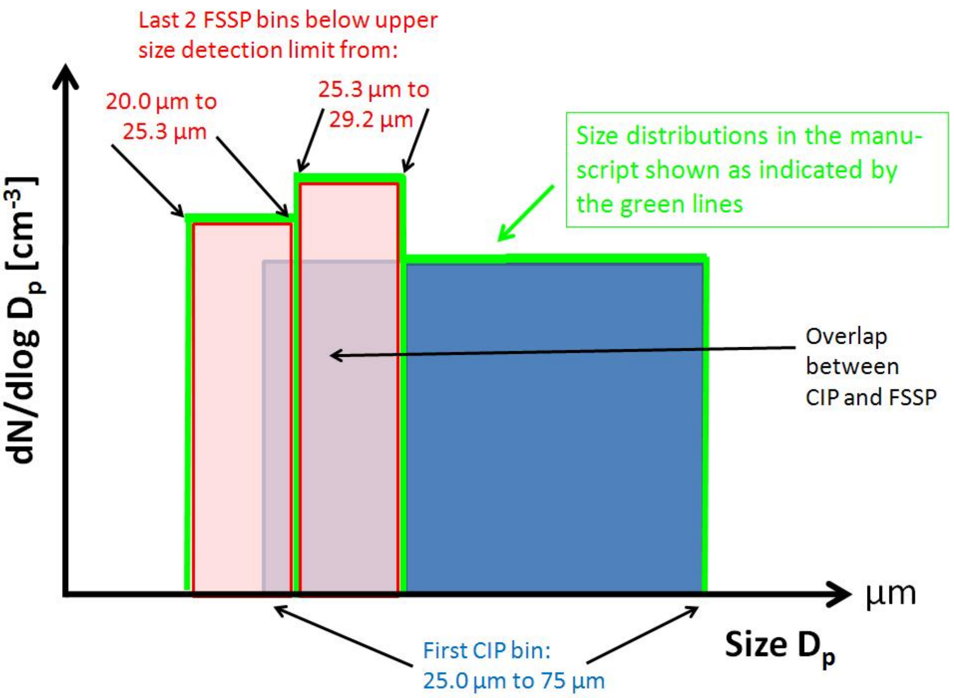
the size distribution of all particles including the shattered particles is shown in black. The blue curve gives the resulting size distribution with removed shattering. The right panel provided the results with linear ordinate for better clarity and one can see that the removal of the shattered particles does not affect the size distribution significantly. For the most part the difference probably is within the limits of uncertainty due to counting statistics and the sample volume. However, if requested we can additionally provide the red curves for the corresponding figures in the paper.

3. General treatment of FSSP data and relation to shattering artefacts:

3.1. Sizing of FSSP data into bins: Application of the T-matrix method: In principle the size bins from the T-matrix model of the FSSP (models 300 and 100) scattering geometries after Borrmann et al., (2000; Application of the T-matrix method to the measurement of aspherical (ellipsoidal) particles with forward scattering optical particle counters, J. Aerosol Sci., 31, 789-799) can be used. However, the adopted T-matrix code does not converge anymore for particles with sizes above 16µm diameter.

The FSSP data we used in the manuscript extend from 2.7µm to 29.2µm. Based on the T-matrix method this range could be subdivided into 15 bins including three bins from 16µm to 29.2µm. It needs to be assumed that the T-matrix results can be extrapolated to these last 3 bins. In practice we subdivided the size range from 2.7µm to 29.2µm covered by the FSSP here *only into 7 bins*. This artificial reduction of the size resolution was done by carefully inspecting the corresponding scattering cross sections from the Mie- and the T-matrix curves and defining the bin limits “manually”. The reason is, that in fact it *is* difficult to apply the FSSP for ice particles and that the T-matrix method only can serve to demonstrate –within narrow limits- that it is *not impossible* to measure inside cirrus. (This was the original intent of the Borrmann et al., 2000, paper.) To be conservative and reduce potential cross-sensitivity, where particles are counted into bins they do not belong to, we decreased the size resolution from 15 to 7 bins. For the subvisual cirrus data the counting statistics mostly is not good, such that a further reduction of the size resolution to only 6 bins is justified even more. In summary, responding to the reviewers questions: Yes, we used the T-matrix scattering cross section curves but decreased the size resolution to one half of the theoretically possible number of bins.

3.2. Combined FSSP and CIP size distributions: As shown in the drawing below the size distributions from the FSSP are simply overlaid onto the CIP size distributions. The first CIP size bin actually extends down to sizes of 25µm but is only displayed down to 29.2µm which is the upper size limit of the FSSP size distribution. In the manuscript the presented size distributions consist of what is delineated by the green lines. No smoothing or running averages have been applied to the size distribution data and the unaltered measurement results are shown with error bars based on propagation of counting statistics and sample volume uncertainties. For us this constitutes the most transparent and honest approach of presenting the data. The size distributions from the FSSP are considered as contaminated by shattering in *all* bins, if the two highest bins (shown in red above) do not well overlap with the lowest CIP bin (shown in blue).



3.3. Rejection of FSSP data based on poor overlap with CIP: Such cases of poor overlap always exhibit much higher FSSP concentrations in the last two bins than the CIP in its first bin. We consider the CIP as reporting the correct concentrations (after application of the state-of-the-art corrections), because unlike for the FSSP we here have tools like interarrival time analyses available for identification of shattering events. If both instruments exhibit overlap within their error bars, we considered the FSSP data not to be significantly influenced by shattering.

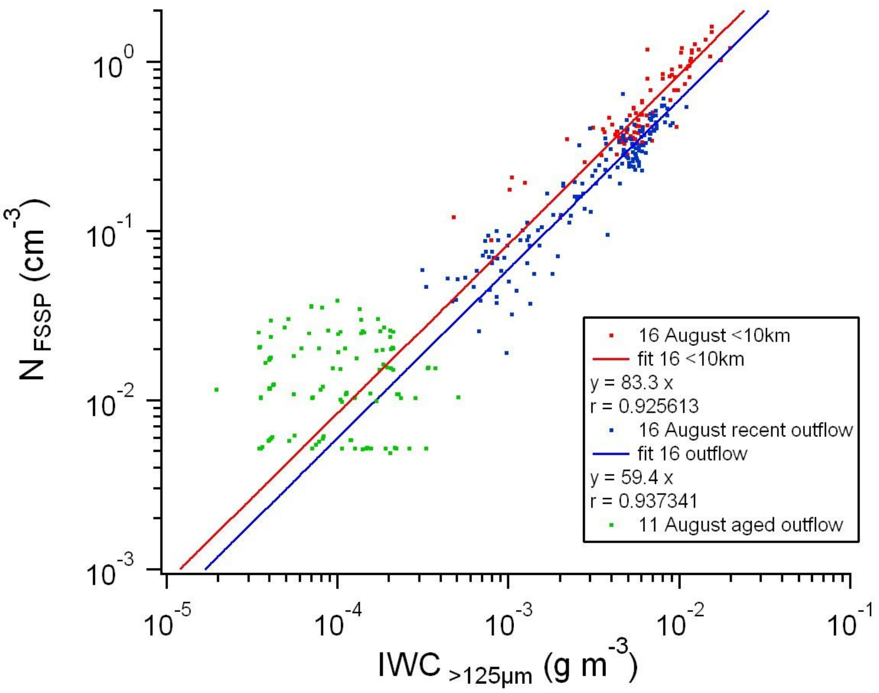
In reply to the point made by referee Grant Allen it turned out that the volume of FSSP (and CIP) data that had to be rejected based on poor overlap is rather small. All relevant measurements for this study were obtained at potential temperatures above 345K and poor overlaps were mostly found below this level. Data from below roughly 10 km are not shown in the paper except for the time series of 16 August in Figure 13 and size distributions in Figure 14. From the measurements above 345K, only 3 size distributions needed to be discarded. Specifically, on the flight on 8 August one and on the flight on 11 August only two size distributions (accumulated from time periods of 10 and 20 seconds, respectively) were removed at potential temperatures between 345K and 346K. Thus, only the lowest potential temperature bin in Figure 1 (lowest panel) of the manuscript is affected from such rejection. Data obtained in the outflow regions on 7 August and 16 August and in the subvisual cirrus cases are not affected by data rejection due to poor overlap.

Furthermore, due to a problem in the data acquisition software (which was easily recognised a-posteriori) some data shortly before and during the MCS outflow crossing on 7 August was lost. This lost data should also have been measured at the lowest considered potential temperatures (345K - 350K). The measurements obtained in the outflow regions on 16 August and 11 August and in the subvisual cirrus cases are not affected by this problem.

“Poor overlap conditions” seem to occur preferably when the Geophysica performs particular manoeuvres like narrow turns. It has to be borne in mind that good overlaps between aircraft borne aerosol instruments of such different nature are all but common occurrence. For example the agreement in the overlapping size ranges of the FCAS and FSSP-300 instruments deployed on the ER-2 in conditions, which are simple by comparison, namely within the stratospheric Pinatubo aerosol, was much worse (Wilson et al., In situ observations of aerosol and chlorine monoxide after the 1991 eruption of Mount Pinatubo: Effect of reactions on sulphate aerosol, Science, 261, 1140-1143, 1993; Jonsson et al., Performance of the Focused Cavity Aerosol Spectrometer for measurements in the stratosphere of particle-size in the 0.06 - 2μm diameter range, JAOT, 12, 115-129, 1995).

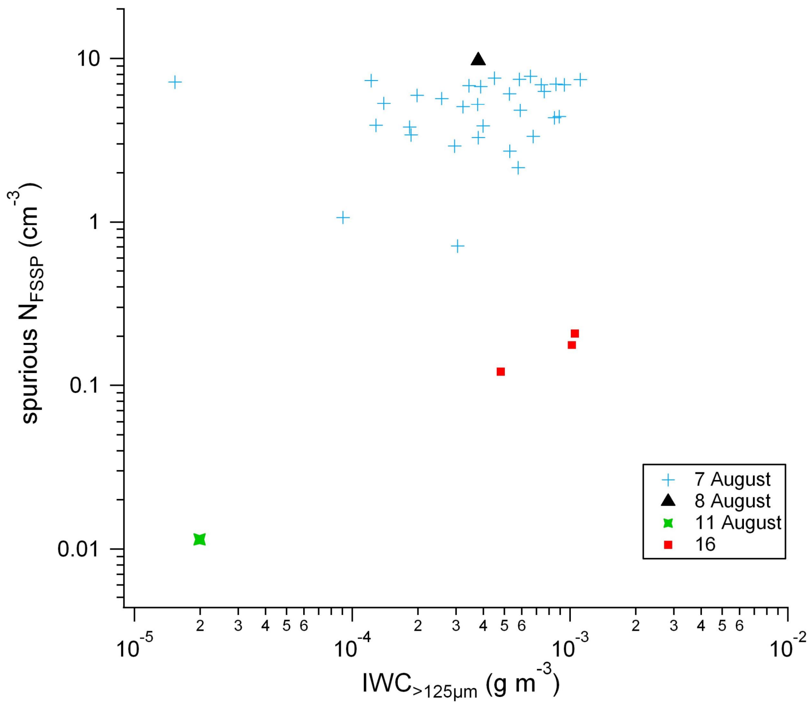
3.4. Correlation analysis following the Jensen et al., 2009, paper: Reviewer #4 suggested to perform similar analyses as in Jensen et al. (2009, On the importance of small ice crystals in tropical anvil cirrus, Atmos. Chem. Phys., 9, 5519–5537) using correlations between large particle IWC and small particle number concentrations in order to identify measurements potentially affected by shattering. We conducted such analyses and report the results in this section.

The figure below shows the concentration of FSSP particles vs. IWC from the CIP particles above 125μm for the flights from 11 and 16 August. The data from 11 August are from inside and the vicinity of aged outflow events, and show no correlation. This follows the expectation of Jensen et al (2009) for *aged* clouds in case shattering is not significant. By contrast, Jensen et al. (2009) do expect a correlation for *young* ice clouds based on microphysical arguments. Furthermore, they anticipate such a correlation in case shattering introduces significant amounts of artefacts. Indeed, the measurements from 16 August in young clouds and recent outflow also exhibit such a correlation. We see more shattering for the cloud measurements below 10km from the CIP data, and the red data points (plus line) indicate a stronger slope of the correlation compared to the recent outflow case in blue data points and line (Note the log scales and the absolute difference in the coefficients for the slopes). The young cloud data of 7 August show a similar figure as those from the 16 August but exhibit a somewhat worse correlation.



Thus, our analysis confirms the expectations of Jensen et al. (2009).

In order to quantify possible shattering Jensen et al. (2009) suggested a further step for analysis. They applied 5 different filter criteria to their CAS data to find enhanced, spurious concentrations due to shattering.



As can be seen from the figure above there are only very few measurements left in our data after this filtering is applied. For 7 August, only, enough data points (light blue) remain after the filtering to perform a correlation analysis. However, only a poor correlation results with a Pearson coefficient of 0.05, which indicates that the few spurious particles do not depend on IWC, i.e. "plus-minus" the presence of the largest particles. Thus, shattering seems to have not impacted our measurements to a significant amount.

In summary our data confirm the assumption of Jensen et al., 2009, according to which (1.) in young outflow scenarios a correlation between the IWC from particles with sizes above 125 µm and the number densities of small particles is present, (2.) no correlation in SVC and aged outflows can be found, and (3.) not as much "alarming" data points survive the filtering as was the case for the CAS.

4. Summary of the effects of shattering in the data of the manuscript:

We summarise here which figures of the manuscript possibly are affected by shattering artefacts and what has been changed in order to highlight potential problems to the reader.

Figure 1: The lowest two panels for the potential temperature bins from 345K to 355K may be influenced by shattering. However, we consider this as insignificant based on the arguments from Section 1.2, 1.4, 3.2, and 3.3. We would like to leave the figure as it is, except for subdividing it in two more potential temperature bins as suggested by Reviewer Grant Allen. We could add a bar into the panel extending parallel to the abscissa from 2.7 µm to 25 µm with the text string "Potential shattering effects not excluded". The panels for the higher potential temperature bins are not affected because there the CIP did not detect particles which are large enough to cause significant shattering, or simply, the CIP did not detect any shattering according to the interarrival time method.

Figure 5 The size distributions "AOF1" and "AOF2" contained less than 10% shattered particles in the CIP size range (1.5% and 7.4%, respectively), furthermore, the overlap of FSSP-100 and CIP size range is excellent. "OF1" and "OF2" both contained 10% shattered particles. Therefore, we highlight the size range of the FSSP in these size distributions and include a remark in the caption “size range potentially affected by shattering”. For the lowest panel we proceed as in Figure 1.

Figures 10 and 16: These figures are not affected by shattering because the largest reported particles are too small and the interarrival times for the CIP were by far above the threshold value.

Figure 14: The mean percentage of shattered particles in the CIP size range for the cases “BOF1”, “BOF2”, “OF3”, and “OF4” is above 10% (12.8%, 14.5%, 12.2%, and 13.5%, respectively). Therefore, the FSSP part of these size distributions is highlighted as in Figure 5. “OF5” contains less than 10% shattered particles (8.4%), furthermore, the number densities of particles larger than 500 µm is very low and the analyses shown in Section 3.4 indicates that this event is not affected significantly by shattering. We can do highlighting as in Figures 1 and 5.

Figure 19: Here, only the upper two red curves might be influenced from shattering in the size range below 30µm, although we believe this is not significant based on the arguments in Section 1.2, 1.4, 3.2, and 3.3. We can remove the data of the FSSP from these two curves or replace them with dashed lines and a corresponding remark in the caption. The black curves are most likely not affected because of the low number concentrations of the large particles and the interarrival times of particles detected by the CIP were by far above the threshold value. The green curves may be contaminated by shattering in the FSSP size range (although we believe this is insignificant) while the data for the CIP have been corrected for shattering using the interarrival times.

We believe with these measures we have performed state of the art analyses as far as these are possible for the adopted instruments and we provided best possible transparency on the use of the data to enable readers to arrive at their own opinions. We also believe that the data should be presented -provided all these caveats highlighted properly- because such measurements from West Africa are practically non-existent, and further research like this will be next to impossible in the near to mid range future. Finally, we suggest including these statements amounting to “a paper within the paper” in the supporting online material of our manuscript for the interested reader. In this case we would put a reference to the supporting material in the revised version of the manuscript in Section 3.1.1 (titled “Shattering of ice particles on the cloud particle probes”).