Author responses to Reviewer 1

Format: The reviewers' comments are in black font while author responses are in red font. Text in red font italics indicates revised/added text in the revised manuscript.

We understand that reviewing this paper took a lot of time and effort, and we sincerely thank you for your comments that have improved this paper. Below are our responses to the general and specific comments:

The manuscript by Mitchell et al. improves upon the development of an earlier methodology published in 2018 to retrieve the layer properties of optically thin cirrus clouds using a combination of the IIR and CALIOP instruments on the CALIPSO satellite. The synergistic combination of lidar and infrared measurements is a unique property of the CALIPSO satellite and were initially conceived for the very purpose of retrieving cirrus microphysics. The proposed algorithm developed in this manuscript seeks to use the ratio of the 12 micron to 10 micron radiances to derive a relationships between absorption optical depth and microphysics and the vertically integrated attenuated lidar backscatter that is related to the visible optical depth at 523nm to derive the layer-mean ice water content, the effective particle size, and ice crystal number concentration, Ni. The retrieval is formulated using a closed set of simplified equations for the radiative properties. The closure is achieved using analysis of in situ cirrus particle size distribution data from multiple cirrus measurement campaigns along with assumptions regarding the absorption properties of the PSD. The authors compare statistical results of the retrieval algorithm (i.e. regionally derived cirrus properties) to the statistics of the cirrus properties in the campaigns. Overall, I find the methodology unique and innovative builds on a long history of cirrus data analysis by the authors. However, I have several major concerns that I think the authors must address in a major revision of the present manuscript.

1. The algorithm is dependent upon in situ cirrus data particle size distributions for relationships that allow for closure of the algorithm equations. While the general approach seems sound, there are several critical issues that should be addressed in more detail. The first of these is the fact that the algorithm retrieves the properties of a layer that may be several kilometers deep where vertical variations in cirrus properties are often considerable over the depth of the layer. However, aircraft data are typically collected along level flight lines within cirrus layers and they do not typically sample the vertical structure. Assuming that the relationships developed from the in situ data are relevant in small volumes, why would the relationships apply to entire cirrus layers particularly when the infrared absorption properties are likely weighted to a different region of the layer than the visible optical depth?

   The relationships derived from in situ data relate $\beta_{eff}12/10$ to N_i/A_{PSD}, N_i/IWC, and $1/Q_{abs,eff}$ (12 µm). N_i/A_{PSD} is the inverse of an area (units are cm^{-2} in Fig. 14), N_i/IWC is the inverse of a mass (units are g^{-1} in Fig. 14). $1/Q_{abs,eff}$ (12 µm) has no unit and $\beta_{eff}12/10$ is a ratio of effective absorption optical depths. There is no notion of volume involved, which justifies applying the relationships to layers.

   We show in Sect. 2.2.5 how we use the CALIOP extinction profile in the visible to estimate the IIR weighting function. In panel a where the cloud emissivity at 12.05 µm is small (0.06), the IIR weighting function and the CALIOP extinction profiles have very similar shape. In panel b, cloud emissivity is larger (0.44) and we see that relative

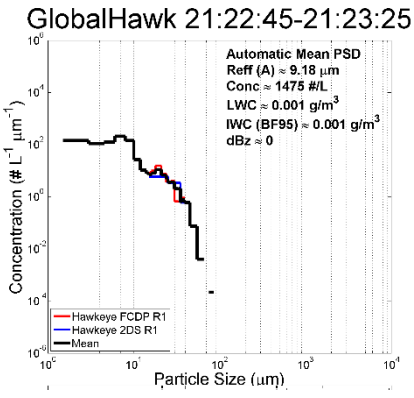
to the CALIOP extinction profile, there is more weight at 12.05 µm in the upper part of the cloud and less in the lower part. The inferred IIR equivalent layer thickness represents the portion of the cloud contributing to the absorption optical depth. The cloud properties such as extinction (α_{ext}), N_i or IWC characterize the layer.

We corrected and clarified the beginning of Section 2.1 (that briefly describes the M2018 retrieval) as follows (changes in italic):

« *Layer* N_i is derived from the N_i/IWC ratio after retrieving *layer* IWC from α_{ext} and the empirical D_e - β_{eff} relationship. The uncertainty in N_i can be reduced by eliminating its dependence on the empirical D_e - β_{eff} relationship and by replacing the N_i/IWC ratio with the N_i/A_{PSD} ratio, where A_{PSD} is the PSD projected area *per unit volume* directly measured by the 2D-S probe (Lawson et al, 2006; Lawson, 2011). »

2. Continuing with the previous comment, the PSDs measured in cirrus by probes are well known to contain shattering artifacts.While all modern probes now contain tips that reduce shattering and data are analyzed to remove clusters of small particles that likely resulted in shattering, this effect is still present and likely dominates (or at least biases) the number concentration of ice crystals, Ni, in aircraft measurements. I note that Kramer et al. (2020) in their cirrus climatology spend a large section of their analysis on this topic and still feel that it biases their results. The PSDs shown in Figure 8 in this manuscript are representative where there is a main mode of larger particles that is in the several hundred-micron size range and then tails off smoothly toward smaller sizes (like a modified gamma distribution would do) but this mode is interrupted in the 10's of microns range by a smaller mode that occupies the smallest several size bins. Can the authors explain the microphysical mechanism that would result in this behavior? In my experience, this mode of small particles is omnipresent in cirrus measurements, and, in my opinion, it likely represents shattering artifacts that have not been mitigated by tips or removed through software. I'd be happy if the authors could explain to me why I am wrong about this. Can the authors point to in situ data that does not show this small mode? However, it remains the case that this residual small mode biases the Ni in their data so the presence of the small mode should be addressed either via explaining the microphysical processes that produce it or by finding some way of removing it that is more sophisticated than simply dividing it by 10.4 as described in the paper that they apply to the first bin of the 2DS data.

While the PSDs shown in Fig. 8 are typical, there are many examples where the PSD small mode (peaking ~ 2 microns) is absent (i.e., the concentrations of small crystals with size exhibit a “flat” behavior). And there are other examples where the PSD tends to be unimodal, without a distinct small mode. In these cases, Ni tends to be anomalously high, suggesting homogeneous freezing nucleation (hom) dominates for such PSDs. An example from the ATTREX campaign at -80 °C is shown below:



Here it is seen that ice particle sizes are less than 60 μm, making shattering very unlikely due to the blocky, simple shapes of the crystals and their sizes. When the PSD is unimodal and Ni exceeds 200 L^{-1} (Ni = 1475 L^{-1} in this case), this is typical of hom. Hence, hom can be an abundant source of small ice crystals and may contribute to the PSD small mode. That is, a pronounced small mode is not necessarily an indication of ice artifacts from shattering. The details of how turbulence-induced transient hom events affect the PSD are addressed in Kärcher et al. 2025, https://doi.org/10.1038/s41612-025-01024-w.

3. The authors address algorithm uncertainty in an appendix by deriving a number of equations that propagate the error in the measurements through to the retrieved quantities. While this seems thorough enough, they do not present actual results of the uncertainty analysis – just the equations. They do mention that pixel to pixel uncertainty in Ni can be on the order of a factor of 2, they do not, however, present an actual analysis of the error or show how it depends upon the uncertainties in the input variables or assumptions. Such an analysis is a requirement in my opinion.

As noted by the reviewer, the uncertainties are estimated by propagating the error in the IIR measurements to the retrieved quantities. Uncertainties in the X-β_{eff} relationships are not included. This was clarified by adding the following sentence in Sect. 3.4 :

« *Note that additional uncertainties in the X- β_{eff} relationships are difficult to estimate and are not included in this assessment.* ».

An analysis of errors is provided in Table 6 for the ATTREX campaign and in Table S1 in the Supplement for the POSIDON campaign. These uncertainties are discussed in Sect. 4.1, as shown below :

« Table 6 shows median retrieved properties and relative uncertainty estimates for ATTREX for cirrus having ~ 0.2 - 0.3 < τ < ~ 3 (left) and ~ 0.01 < τ < ~ 3 (right). A similar table for the POSIDON campaign is shown in the Supplement (Table S1). In Table 6, median $\Delta\beta_{eff}$ ranges from 0.03 to 0.44 where median τ = 0.05 at T_r = 193 K, $\Delta Ni/Ni$ = 1.88 and $\Delta D_e/D_e$ = 0.98. The smallest median $\Delta N_i/N_i$ is 0.35 at T_r = 193 K when only the thicker clouds are sampled and median τ is somewhat small (0.23), but this is compensated by the fact that β_{eff} = 1.57 where the sensitivity of the technique is very favorable. In contrast, median β_{eff} is 1.056-1.058 at 233 K, where the sensitivity of the technique is less favorable, which explains the occurrence of relative uncertainties larger than 2.4 despite the small $\Delta\beta_{eff}$ = 0.03.

Again, these uncertainty estimates characterize random uncertainties of individual retrievals and are reduced for statistical analyses involving a large number of samples. »

Uncertainties are shown and discussed in the case study which is now added at the end of Sect. 4.2 (see comment #4).

4. The authors move from algorithm development directly to statistical comparisons with other campaigns. They do not show actual data collected by the sensors and the retrieval results with error bars that would result from application of their algorithm to actual cirrus layers. Furthermore, it seems necessary to demonstrate an actual case or two in circumstances when in situ aircraft data were being collected data underneath the satellites. The SPartICus data that is used by the authors in their analysis has approximately two dozen flights where the Lear Jet flew along the paths of the CloudSat and CALIPSO satellites. These are well documented in Deng et al. (2013; *DOI: 10.1175/JAMC-D-12-054.1*). As a matter of fact, since the authors use SPartICus data, it seems reasonable for them to replicate the results in Deng et al. This would be quite straightforward. The authors also use data collected in TC4. Mace et al. (2010; *doi:10.1029/2009JD012517*) illustrate a case when the NASA DC8 flew along the CloudSat and CALIPSO tracks in tropical cirrus. Replicating these direct comparisons between algorithm results and in situ aircraft data is a necessary step toward establishing confidence in the algorithm results and would go a long way toward addressing many of the earlier concerns I have raised.

We move from algorithm development to statistical comparisons with in situ observations because comparing layer retrievals and in situ observations is not straightforward, but we agree that this should be illustrated.

We added a case study during the SPARTICUS campaign at the end of Sect. 4.2 for 30 March 2010. The case studies shown in Deng et al. (2013) could not be used because the scenes were not adpated for our IIR retrievals for the following reasons. The cirrus is prevailingly opaque to CALIOP for the thick-anvil case on 17 April 2010 and the thick-cirrus case on 12 June 2010. The thin-cirrus case on 22 April 2010 could not be used because low clouds were present below 5-km altitude. For the 1 April 2010 case, a portion of the cloud is a single-layer semi-transparent ice cloud but the radiative temperature is most of the time warmer than 235 K.

Below is the case study and the new text added at the end of Sect. 4.2 :

The difficulty to directly compare IIR layer retrievals and aircraft in situ data is illustrated in the SPARTICUS case study shown in Fig. 21 for 30 March 2010. Following the CALIPSO track, the Learjet flew northwards (leg 1, triangles) with measurements at 11 km altitude 7 to 3 minutes before the CALIPSO overpass and then southwards (leg 2, diamonds) with measurements at 11.6 km altitude 6.5 to 8 minutes after. CALIPSO detected a single layer cirrus of top altitude near 12.6 km. The colors in panel a represent the altitude-dependent CALIOP extinction profiles scaled to IIR τ. The colors inside the triangles and diamonds indicate the PSD extinctions larger than 0.01 km^{-1} after averaging over a 30-s period. At the top of panel a is IIR layer α_{ext}, which was derived from τ and Δz_{eq} shown in panel b. As discussed in Sect. 2.2.5, Δz_{eq} represents the portion of a layer contributing the most to the cloud emissivity. The solid black line in panel a is the radiative altitude corresponding to T_r, which to a first approximation corresponds to the mid cloud altitude (see Fig. 7). IIR D_e in red in panel c (with vertical bars indicating $D_e \pm \Delta D_e$) is lower than the in situ values, which is explained by the fact that both legs were below the radiative altitude. That is, in the lower half of an ice cloud, mean ice particle size tends to be larger and N_i lower relative to the

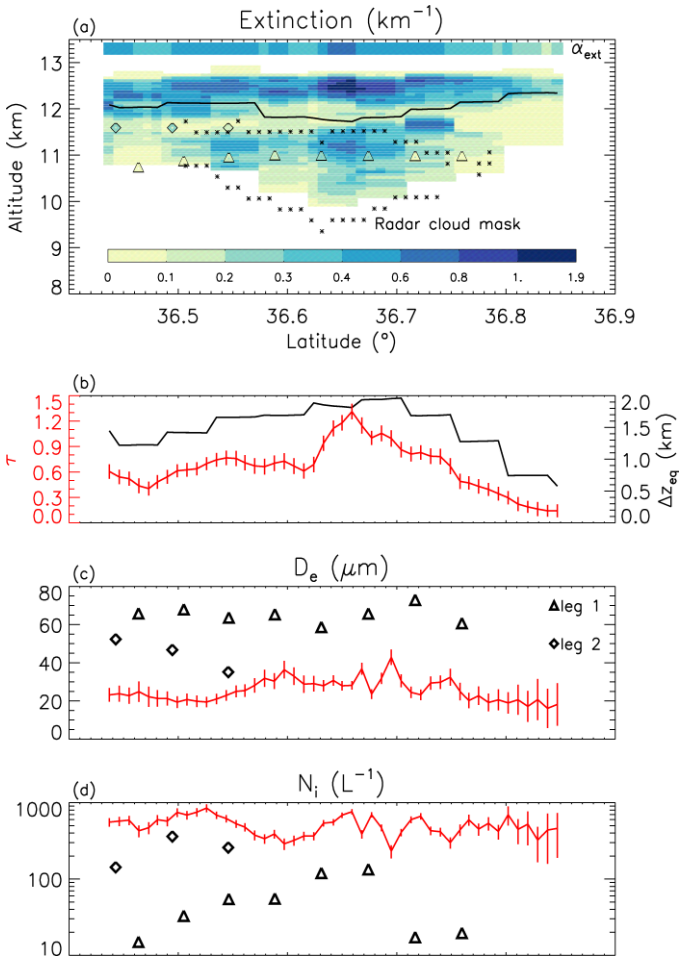


Figure 21. Comparison of IIR retrievals and in situ observations on 30 March 2010 during the SPARTICUS field campaign (CALIPSO granule 2010-03-30T19-27-25ZD). (a) extinction profile derived from the CALIOP lidar, IIR layer α_{ext}, and PSD extinctions in leg 1 (triangles) and leg 2 (diamonds). The stars denote the boundaries of the CloudSat radar GEOPROF cloud mask; (b) IIR τ (red) and Δz_{eq} (black, right axis); (c) IIR (red) and in situ (triangles and diamonds) D_e; (d) same as (c) but for N_i. The vertical bars in red in panels b-d represent the IIR estimated uncertainties.

upper half due to diffusional growth and aggregation (e.g., Mitchell, 1988; 1994; Field and Heymsfield, 2003). Only a lower portion of the cloud was detected by the CloudSat radar (shown by the stars) between latitudes 36.5° and 36.78°. IIR D_e is the smallest (around 20 µm) south of 36.5° and north of 36.78° where there is no radar detection, indicating crystals smaller than about 40 µm. Moreover, the absence of radar detection outside this CloudSat domain (defined by the stars) indicates ice particles smaller than ~ 40 µm, revealing a vertical gradient in ice particle size. Regarding N_i (panel d showing $N_i \pm \Delta N_i$), the large IIR N_i values in red between 300 and 850 L^{-1} are explained by large N_i near cloud top. Regarding uncertainties, ΔN_i is overall equal to about 80 L^{-1} and its noticeable increase up to 300 L^{-1} in the northernmost part of the cloud is due to the decrease of τ. The same observation applies to ΔD_e which is between 3 and 11 µm. To summarize, while the vertically resolved extinction retrievals exhibit reasonable agreement with the in situ extinction measurements, the bulk cloud layer retrievals often do not exhibit similar agreement, and this appears to be due to vertical gradients in D_e and N_i and aircraft sampling location. This case study has been classified as ridge crest cirrus which have higher N_i than the other cirrus cloud classes described in Muhlbauer et al. (2014). In this regard, the retrievals here are consistent with this category of cirrus cloud.