

RESPONSES TO THE REFEREE #1 COMMENTS

We thank the referee for his/her review of this manuscript, and for constructive comments that have significantly improved this paper. We understand that this is a serious undertaking that requires considerable time and effort, and your efforts are appreciated!

In the text below, black font is used for the referee comments (RC) and blue font is used for author comments (AC) and new text added to the paper.

General comments:

RC G1. A major concern with the current manuscript is its length. The authors cover a very wide range of information and results that are all interesting but their amount is, in my opinion, counterproductive. It should be kept in mind that this study will be of high interest to diverse communities, as it contains interesting aspects for remote sensing, in situ measurements, modeling or ice microphysics. Based on the current version I suspect that readers will only focus on their section of interest and miss important results. One idea could be to shorten section 2, since a big part of it has already been described in Mitchell et al. [2016]. The results related to the $N(D)_{1=0}$ analyses are very important and should be mentioned in the paper but several of them could be moved to supplementary materials in order to lighten the discussions and the figures. Other analyses that are not focused on N (the novelty of the paper), such as section 5.4, could also be moved to supplementary materials. This would allow to remove some technical discussions and focus the paper on result analyses, which are also better suited for a publication in ACP.

AC G1. Both referees commented on the paper's length being too long, although the content was appropriate. To remedy this problem, we moved 11 of the original figures and two new figures to a new file titled "Supplementary Materials". Some of the original text was also moved to Supplementary Materials. This shortened the paper's length by about 8 pages.

RC G2. The authors acknowledge in the manuscript that this new retrieval method is not expected to provide accurate absolute values of N. Only the spatial variability of this parameter is well represented. The accuracy of the absolute N is discussed throughout sections 3 and 4 but would it be possible to summarize a final estimate of the uncertainties on N in the conclusion? By this I mean not only the $\Delta N/N$ but also combining what has been learned from the in situ evaluation. For instance, would it be a factor of 2, 5, one order of magnitude? Also, in what conditions can optimal retrievals be expected? This type of information would be very useful to future users of this dataset, especially since this paper will serve as a reference.

AC G2. In the 1st paragraph of the new Sect. 3.2 (previously Sect. 3.3), we write that ΔN is computed assuming a negligible error in the relationships.

We added the following sentence to clarify that: *“Errors in the relationships create additional systematic uncertainties, as discussed in Sect. 3.1.”*

As suggested by the reviewer, the following text is added in the 2nd paragraph of Summary and Conclusions:

“Four formulations of the retrieval were used, based on either the SPARTICUS or TC4 field campaigns, with the smallest size-bin of the 2D-S probe either assumed valid ($N(D)_1$ unmodified) or by assuming $N(D)_1=0$. The SPARTICUS unmodified $N(D)_1$ assumption gives the highest N values while the TC4 $N(D)_1 = 0$ assumption yielded the lowest N values. The N predicted from these two formulations differed by a factor of two (see Fig. 5), thus defining a possible systematic uncertainty in N . The random relative uncertainty, $\Delta N/N$, associated to a N retrieval is most of the time < 0.5 for small crystals ($D_e < 35\text{-}50 \mu\text{m}$), but increases up to more than 2 at the sensitivity limit”.

In addition, the definition of β_{eff} is now repeated at the beginning of the conclusion and the sensitivity ranges of the retrievals is summarized at the end of the 1st paragraph of the conclusion as:

“Perhaps the most unique aspect of this retrieval method is its sensitivity to small ice crystals via β_{eff} . The sensitivity ranges ($N/IWC > \sim 10^7 \text{ g}^{-1}$ & $D_e < 90\text{-}110 \mu\text{m}$) are usually compatible with cirrus clouds ($T < -38 \text{ }^\circ\text{C}$) since PSDs tend to be narrower at these temperatures, containing relatively small ice particles.”

Finally, we note at the end of the 4th paragraph under Summary and Conclusions: “When the sample selection criteria were relaxed to accept samples having cloud OD > 0.1 (instead of OD > 0.3), median N was reduced by a factor of 1.5 on average while $\Delta N/N$ was more than doubled.” This is also mentioned earlier in the paper at the end of Sect. 5.2.

RC G3. A clear limitation of the method is its necessity to filter out a lot of data based on different cloud conditions. This, as shown in table 4, leads to samples that are only representative of up to a 2% frequency of occurrence. Is there a particular reason for not using more data?

AC G3. The rationale for selecting the relevant cloudy scenes for this study is presented in Sect. 2.2.1. We tried to clarify, and Sect. 2.2.1 now reads (changes are in bold):

“Because IIR is a passive instrument, meaningful retrievals are possible for well identified scenes. This study is restricted to the cases where the atmospheric column contains one cirrus cloud layer. We also insure that the background radiance is only due to the surface (see Eq.

(3) allowing a more accurate computation than for cloudy scenes. The retrievals were applied only to single-layered semi-transparent cirrus clouds that do not fully attenuate the CALIOP laser beam, so that the cloud base is detected by the lidar. The cloud base is in the troposphere and its temperature is required to be colder than -38°C (235 K) to ensure that the cloud is entirely composed of ice. This is likely to exclude liquid-origin cirrus clouds from our data set (Luebke et al., 2016). **When the column contains also a dense water cloud, the background radiance can be computed assuming that the water cloud is a blackbody. However, because systematic biases were made evident (Garnier et al., 2012), we chose to discard these cases, which reduces the number of selected samples by about 25 %.** Because the relative uncertainties in τ_{obs} and in β_{eff} increase very rapidly as cloud emissivity decreases (Garnier et al., 2013), the lidar layer-integrated attenuated backscatter (IAB) was chosen greater than 0.01 sr^{-1} to avoid very large uncertainties at the smallest visible optical depths (ODs). This resulted in an OD range of about 0.3 to 3.0. Similarly, clouds for which the radiative contrast $R_{\text{BG}} - R_{\text{BB}}$ between the surface and the cloud is less than 20 K in brightness temperature units are discarded. IIR observations must be of good quality according to the quality flag reported in the IIR Level 2 product (Vaughan et al., 2017)."

Frequency of occurrence could be increased by relaxing some selection criteria, but the difficulty is that the additional information could be obscured by large additional uncertainties. Following reviewer #2's suggestion, we reprocessed the data set to include clouds of OD between 0.1 and 0.3. This increased the frequency of occurrence by a factor 2 to 3 (as now shown in Table 4), but this also increased the relative uncertainty in IIR β_{eff} and in CALIPSO N. As noted under AC G2, we have reported how this affects N.

The method seems to be rather straightforward to apply on CALIPSO satellite products, and some of the co-authors should be familiar with such operational treatments. It therefore shouldn't take long to process 10 years of data and greatly improve the statistical significance of the results shown in this study (especially in sections 5 and 6).

The first paragraph of Sect. 2.2 has been shortened and modified to clarify that IIR β_{eff} is based on the CALIPSO IIR Version 3 Level 2 operational product, but that the various improvements implemented for this study required to reprocess the dataset. Processing of other years will be interesting for instance to examine inter-annual variability, and will be carried out with the new version 4 of the operational products. The end of the first paragraph of Sect. 2.2 now reads:

"IIR effective emissivity was reprocessed for this study to reduce possible biases, as described in the following sub-sections. [.....]. These improvements will be implemented in the next version 4 of the IIR Level 2 products. This is why we only focus in this study on a limited dataset as a proof-of-concept. We will further develop a more statistically relevant analysis using the version 4 products."

RC G4. Concerning the results shown in section 4 (comparisons to in situ measurements), more direct comparisons between SPARTICUS measurements and CALIPSO could be obtained by looking at the numerous co-incident overpasses between the satellite and the aircraft. This would avoid issues related to non-representative events in the in situ data by comparison to a long-term regional dataset, as noted in page 22. Have you tried looking at these exact satellite overpasses only?

AC G4. Yes, but unfortunately, we found that spatially and temporally coincident measurements are very rare for scenes meeting the IIR cloud selection criteria. We added the following sentence at the end of the 1st paragraph in Section 4.1:

“Data analysis is performed on a statistical basis, as coincident in situ and satellite data only provide a very small dataset due to our data selection.”

RC G5. Perhaps I have missed this information but I am under the impression that the D_e retrieved by IIR is not used in this method. How different would be the results if this operational retrieval was used instead of the in situ relationship? Is there a reason for not doing this?

AC G5. The fundamental parameter retrieved from IIR is β_{eff} . In the CALIPSO retrieval equation of N (Eq. (8), Sect. 2.5), IIR β_{eff} is related to N/IWC , D_e , and $2/Q_{\text{abs,eff}}$ using regression curves derived from aircraft PSD measurements (Sect. 2.3 and Sect. 2.4). Indeed, the three relationships used in Eq. (8) must be derived in a consistent manner from the same in situ PSDs.

IIR D_e reported in the Version 3 data products is derived using the $\beta_{\text{eff}}-D_e$ relationships presented in Garnier et al. (2013) computed using different optical properties, with no size distribution, and by using both the 12.05/10.6 and 12.05/08.65 pairs of channels. Because the technique and the assumptions are different than in this study, we could not use D_e from the Version 3 product for this study. Furthermore, IIR data have been reprocessed for this study in order to implement the improvement described in Sect. 2.2.2 and 2.2.3, yielding different β_{eff} in this study than in the Version 3 products, especially over land.

We removed the comments about IIR D_e in Sect. 2.2 to avoid confusion and added the following text to the end of the 2nd paragraph of Sect. 2.4:

“It is noted that the $D_e- \beta_{\text{eff}}$ relationships used in the IIR Version 3 operational algorithm (Fig. 3a in Garnier et al., 2013) tend to yield smaller D_e than the relationships derived from these PSD measurements, largely because they were computed with no size distribution.”

Specific comments:

RC S1. Section 2.2: Computing IIR radiances, or converting between optical properties, requires some assumptions on the shape of ice crystals and of the ice particle size distribution.

Dubuisson et al. [2008] for instance showed that the IIR brightness temperatures are sensitive to both parameters. Garnier et al. [2012] showed that converting the IIR effective optical depths into absorption optical depths can also depend on ice particle shape assumptions. Do you have an idea if these assumptions have an impact on your N retrievals? Are they accounted for in the uncertainties described in the appendix?

AC S1. We clarified the presentation of IIR β_{eff} at the beginning of Sect. 2.2, which now reads:

“We use the absorption optical depth $\tau_{\text{abs}}(12.05 \mu\text{m})$ and $\tau_{\text{abs}}(10.6 \mu\text{m})$ retrieved from the effective emissivity in CALIPSO IIR channels 12.05 μm and 10.6 μm . The retrieved optical depths are not purely due to absorption, but also include the effects of scattering. Thus, their ratio is not exactly β , but the effective β or β_{eff} written as

$$\beta_{\text{eff}} = \tau_{\text{abs}}(12.05 \mu\text{m}) / \tau_{\text{abs}}(10.6 \mu\text{m}). \quad (1)''$$

The weak dependence of the retrieval on ice particle shape is described in a new paragraph at the end of Sect. 2.3:

“Cirrus cloud emissivity and τ_{abs} depend on ice particle shape (Mitchell et al., 1996a; Dubuisson et al. 2008). However, this retrieval should not be very sensitive to ice particle shape for several reasons, one being that β_{eff} is directly retrieved from cloud radiances as per (2) and (3). Another reason is that no ice particle shape assumptions are made when calculating β_{eff} from in situ measurements with the exception of the absorption contribution from tunneling (which was not sensitive to realistic shape changes, as described above). That is, the 2D-S probe in situ data include measurements and estimates for ice particle projected area and mass, respectively. MADA optical properties are calculated directly from these in situ area and mass values, thus largely avoiding the need for shape assumptions. Thirdly, this retrieval is most sensitive to the smaller ice particles in a PSD where the variance in ice particle shape is minimal (Baker and Lawson, 2006b; Lawson et al., 2006b; Woods et al., 2018). During the SPARTICUS campaign, many cirrus clouds were sampled so that biases in ice particle shape due to a specific cloud condition are less likely to occur.”

The uncertainties in N described in the appendix are random errors computed assuming no error in the relationships. Errors in the relationships create additional systematic uncertainty (see AC G2), but the dependence on ice crystal shapes should be very weak, as stated above.

Also, are the ice crystal shapes used in the MADA method to compute the in situ relationships consistent with the shapes assumed in the treatment of CALIPSO measurements, and if not is there an impact on the N retrieval and their uncertainties?

CALIPSO IIR β_{eff} is computed directly from the effective emissivity at 12.05 μm and 10.6 μm . This computation does not involve any assumption about the shapes of ice crystals. As such, CALIPSO IIR β_{eff} can be seen as a direct observation. CALIPSO N is then retrieved from CALIPSO IIR β_{eff} and the in situ relationships presented in this manuscript.

RC S2. p. 6 l. 22: By single-layered cloud do you refer to the absence of other ice clouds or also to the absence of a liquid or mixed-phase layer underneath? IIR seems to be able to deal quite well with multi (ice+liquid) layer conditions by adjusting the background radiance. For instance Fig. 10 in Sourdeval et al. [2016] shows that IIR is robust to multi-layer conditions. It would be worth checking if removing this filter makes a big difference in your dataset. Otherwise, not excluding multi-layer scenes would clearly help to increase the global statistical representativity of N retrievals.

AC S2. The rationale for our cloud selection presented in Sect. 2.2.1 has been modified (see AC G3).

We confirm that the retrievals are applied when a cirrus cloud is the only cloud layer in the atmospheric column, that is when the background radiance is from the surface.

We could have also included scenes for which the cirrus cloud layer is over a low opaque water cloud that fully attenuates the CALIOP laser beam. Indeed, in this second case, the background radiance can be computed assuming that the low water cloud behaves as a blackbody emitting at a temperature inferred from the water cloud altitude and GMAO GEOS temperature profiles. The expected accuracy in this second case was evaluated in Garnier et al. (2012) through comparisons of observed and computed blackbody brightness temperatures of low opaque warm clouds. The observations were colder than the blackbody computations by up to 9 K (their Fig. 10), suggesting significant biases in the cloud temperature inferred from the GMAO GEOS profiles in case of strong temperature inversions at the top of the cloud.

We chose to discard these scenes for this study, which reduces the number of samples by about 25 %. These biases will be re-evaluated with the future version 4 of the IIR products which will use MERRA 2 profiles data.

RC S3. Sec. 2.4: Can error bars be added to the points in Fig. 2-4? It seems like these instrumental and computational errors could easily encompass the changes due to the choice of including the first size bin or not.

AC S3. The data points in Figs. 2-4 come from in situ PSD measurements. There is only one PSD per point, and therefore it is not possible to assign an error bar from a statistical basis. One

could in principle propagate errors associated with the various stages of instrument data processing (which perhaps is what the referee is referring to), but uncertainties for these data processing stages have not been characterized in most cases. For example, in “Laboratory and Flight Tests of 2D Imaging Probes: Toward a Better Understanding of Instrument Performance and the Impact on Archived Data” by Gurganus and Lawson (2018, JTech), the 2D-S and the CIP probes are intercompared in various ways, but uncertainties are not estimated that could be applied to address this request. This paper has now been referenced to give the reader additional knowledge of 2D-S performance characteristics. As stated at the end of this paper, “The research reported here does not offer quantitative corrections for the instrument errors that have been uncovered. Instead, it describes conditions under which the uncertainties can exist and provides some qualitative estimates of magnitude. Corrections to the data will require more laboratory tests with upgraded facilities, flight tests, improved forward models of the instruments and numerical simulations.”

RC S4. Fig. 9: Could you comment on the very wide spread (75-25% percentiles) in the Krämer et al N/IWC vs Tc figure by comparison to what is noted for the CALIPSO data? Is there no spread plotted for the Krämer dataset on the right plots or is it too small to be seen? Also please indicate in the caption if the CALIPSO retrievals correspond to the entire 2008 and 2013 periods.

AC S4. To shorten the paper, this figure has been moved to Supplementary Materials and is Fig. S6. The caption states that “in situ measurements from Krämer et al. (2009)” are “shown by the grey curves; top and bottom being minimum and maximum values and middle grey solid curve being the middle value.” Thus, these dashed Kramer curves are not the 25 and 75 percentile curves, but rather indicate the max/min values. To reduce confusion, we have identified curves corresponding to the four formulations as colored. We have indicated that percentiles corresponding to the 4 formulations are not shown in the right panel, and that the retrievals shown are from 2013.

The caption now reads: “Left: Comparisons of the median CALIPSO IIR N/IWC (g^{-1}) for the four formulations (colored) with in situ measurements from Krämer et al. (2009) shown by the grey curves; top and bottom being minimum and maximum values and middle grey solid curve being the middle value. Colored solid curves are median values while dashed curves indicate the 25th and 75th percentile values. Right: Comparisons of CALIPSO IIR β_{eff} shown by the black curves (solid curve gives the median value while dashed curves indicate the 25th and 75th percentile values) with the four (colored) in situ β_{eff} inferred from in situ N/IWC (from Krämer et al. (2009) using the four formulations. Corresponding percentiles are not shown. The navy and light blue curves correspond to the SPARTICUS formulations for the unmodified $N(D)_1$ assumption and the $N(D)_1 = 0$ assumption, respectively. The red and orange curves are using the TC4 formulations for the $N(D)_1$ unmodified and $N(D)_1 = 0$ assumptions, respectively. The CALIPSO IIR retrievals are

from 2013 and are for the approximate latitude range (25 °S to 70 °N) of the in situ data, over oceans (top) and over land (bottom).”

RC S5. Fig. 10e: As a remark, the slightly negative relation between N and T (N decreasing towards low temperature) indicated in the Kramer et al. [2009] study is not found anymore in revised version of the dataset (not yet published but seen in recent conference presentations by M. Krämer et al). This should strengthen the statement p. 31 l. 4 that differences are likely to be due to different cloud sampling.

AC S5. We cannot comment on this since the data is not yet published.

RC S6. Sec. 5.4: Is there an explanation to the fact that Re has a very different dependence on Tc and Tc - Ttop by comparison to what was previously shown for N?

AC S6. In the new Fig. 10 (previously Fig. 13), there is a tendency for N to decrease from left-to-right as a function of $T_c - T_{top}$. In the new Fig. S12 (previously Fig. 16), D_e tends to decrease with decreasing T_c but remains quasi-constant (or less variable) for a given T_c (i.e. less sensitive to $T_c - T_{top}$). D_e is proportional to IWC/A_{PSD} (as calculated for the in situ data and thus implemented in the retrieval regressions), where IWC = ice water content and A_{PSD} = projected area of the ice particle size distribution (PSD). For a gamma function PSD of the form $N(D) = N_0 D^\nu \exp(-\lambda D)$ and representing ice particle projected area as $A = \gamma D^\sigma$ and mass as $m = \alpha D^\beta$, $A_{PSD} = \gamma \Gamma(\sigma + \nu + 1) N / \Gamma(\nu + 1) \lambda^\sigma$ and $IWC = \alpha \Gamma(\beta + \nu + 1) N / \Gamma(\nu + 1) \lambda^\beta$ where Γ denotes the gamma function. The IWC/A_{PSD} ratio now shows that D_e is independent of N, and D_e is proportional to $\lambda^{\sigma-\beta}$ where λ = PSD slope parameter. The difference $\sigma - \beta$ is slightly more than -1 (Mitchell 1996, JAS), showing that D_e is strongly related to the PSD mean size [note that $\lambda = (\nu + 1)/D_{mean}$, where D_{mean} = PSD mean maximum dimension]. The decrease in D_e with decreasing T_c is likely due to a decrease in IWC with decreasing T_c .

New text has been added to the end of the last paragraph in Sect. 5.4 to express these points: *“This weaker dependence on $T_c - T_{top}$ relative to N is expected since D_e is proportional to IWC/A_{PSD} . For a gamma function PSD and assuming power law relationships for ice particle area and mass, D_e is proportional to $\lambda^{\sigma-\beta}$ where λ = PSD slope parameter and σ and β are power law exponents for area and mass, respectively. The difference $\sigma - \beta$ is slightly more than -1, showing that D_e is strongly related to the PSD mean size (inversely proportional to λ). The derivation also shows that D_e is not dependent on N. The decrease in D_e with decreasing T_c is likely due to a decrease in IWC with decreasing T_c .”*

RC S7. Section 6.4: p. 45 l. 11-13: Another possible explanation to the differences in absolute numbers could be that DARDAR-LIM ignores the concentrations of ice particles smaller than 5 μm . p. 45 l. 14-16: As mentioned before, the slight decrease of N towards low T as noted by

Krämer et al. [2009] is not found in the most recent version of their dataset, which is not anymore inconsistent with the relation shown in Gryspeerd et al. [2018]. An increase of N towards low T is also consistent with what is shown in Figures 10ab of this manuscript. Overall, comparisons between the N retrievals presented in this study and DARDAR-LIM are very difficult as both methods are based on different approaches and different instruments. DARDAR-LIM also retrieves vertical profiles of N whereas IIR retrievals correspond to weighted N average values from cloud top. Nevertheless, it is quite remarkable that despite all these differences the two datasets show such similar results. This clearly strengthens the confidence in both satellite products.

AC S7. Section 6.4 has been changed significantly in view of the referee's suggestions. We agree with the referee that absolute differences in N between the DARDAR-LIM and our CALIPSO IIR retrieval may be partly attributed to the fact that DARDAR-LIM (as per the ACP paper) does not account for ice crystals $< 5 \mu\text{m}$ in maximum dimension D. Fortunately, there are DARDAR-LIM results posted on the ACP website as part of the review process for the DARDAR-LIM paper that show N for $D > 1 \mu\text{m}$. We have cited this work to show that using a $D > 1 \mu\text{m}$ cutoff increases DARDAR-LIM N by a factor of about 1.7, making the CALIPSO IIR N retrievals in general agreement with the DARDAR-LIM N values. New text has been added accordingly:

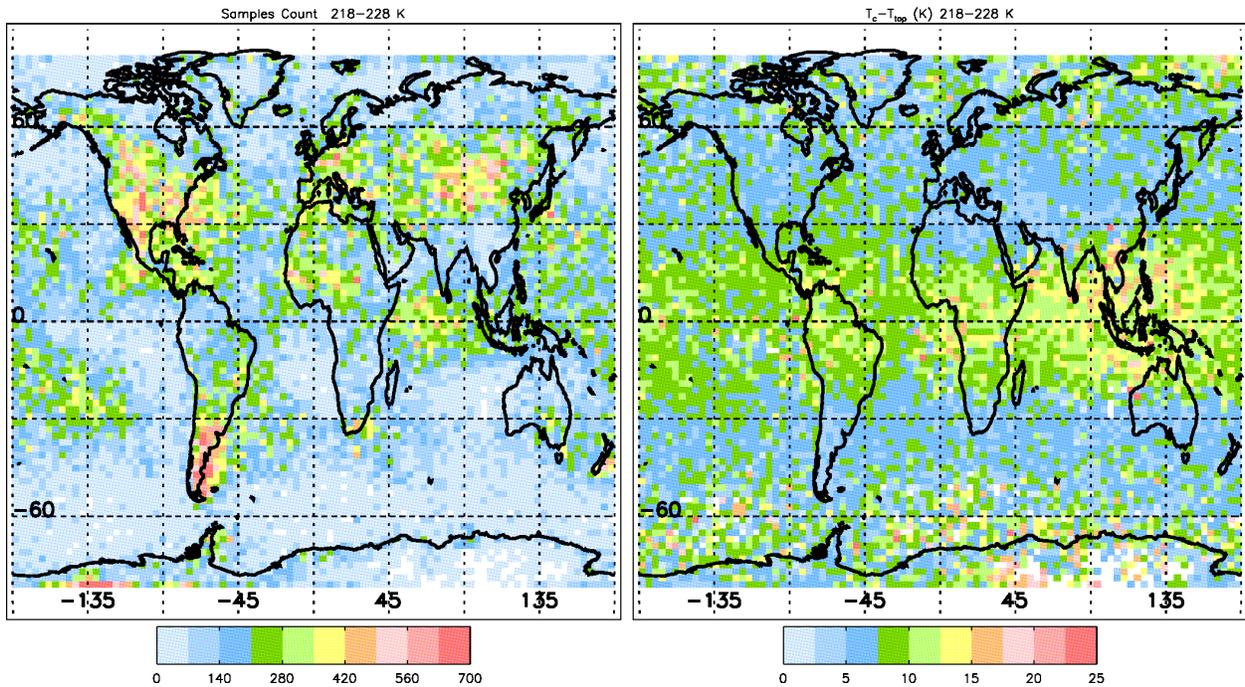
"The highest N values reported in Sourdeval et al. (2018a) and Gryspeerd et al. (2018) are for ice particles larger than $5 \mu\text{m}$. These values would be higher by a factor of about 1.7 if all ice particles larger than $1 \mu\text{m}$ were considered (Sourdeval et al., 2018b). This is important to note since the CALIPSO IIR retrieval has no cut-off size, and it is very sensitive to ice crystals in the 1 to $5 \mu\text{m}$ range (Fig. 1). Keeping this in mind, the CALIPSO N retrievals are comparable in value to the DARDAR-LIM N retrievals."

Regarding the temperature dependence of N, although we cannot cite unpublished work, we appreciate this information and have removed the discussion about the temperature dependence of in situ N measurements (in relation to retrievals).

We agree that there is good qualitative agreement between our CALIPSO IIR and the DARDAR-LIM retrievals, and that this "strengthens the confidence in both satellite products."

RC S8. Fig. 17: It would be interesting to see the spatial distribution of the frequency of occurrence of retrievals and of the distance from cloud top (in terms of temperature) corresponding to this figure.

AC S8. Maps of the number of samples (left) and of Median $T_c - T_{\text{top}}$ (right) associated with the new Fig. 12 (previously Fig. 17), with T_c between 218 K and 228 K, are shown below:



These figures have been added to the Supplementary Materials section of the manuscript as Fig. S13. New text has been added in the 1st paragraph in Sect. 6.4:

“The geographical distributions of the number of samples and of median $T_c - T_{top}$ corresponding to Fig. 12 are shown in Fig. S13 under Supplementary Materials. Median $T_c - T_{top}$ is up to 20 °C and is smaller at mid- to high latitudes than in the tropics.”

The 2nd paragraph in Sect. 6.4 starts now as: *“Results in Fig. 12 can be compared with the retrieved N values near cloud top at 223 K shown in Fig.1 of Gryspeerd et al. (2018).”*

The reviewer’s question led us to comment briefly on the typical relationship between $T_c - T_{top}$ and $T_{base} - T_{top}$. Thus, isolines of $T_{base} - T_{top}$ have been added to Fig. 10 and Fig. S11, and the text has been modified accordingly in Sect. 5.3 (changes in bold):

*“Our retrievals are now examined against both T_c and $T_c - T_{top}$ to estimate the impact of the distance from cloud top. **These N retrievals are shown in Fig. 10 using the SPARTICUS $N(D)_1$ unmodified assumption in the tropics (0-30°) and at mid- (30-60°) and high (60-82°) latitudes in the winter and summer seasons (using both hemispheres) by distinguishing retrievals over oceans and over land. The associated number of samples is given in Fig. S11. Added to these figures are isolines of temperature differences between cloud base and cloud top ($T_{base} - T_{top}$). For most of the layers, $T_c - T_{top}$ represents 30 to 70% of $T_{base} - T_{top}$ (Fig. S11). Fig. 10 shows a strong dependence of N on $T_c - T_{top}$, with large N ($> 500 L^{-1}$) seen near the top of the geometrically thin clouds ($T_{base} - T_{top} < 15$ °C), when $T_c - T_{top}$ is smaller than about 5 °C.”***

RC S9. p. 49 l. 9-10: The "four formulations" have not yet been mentioned in the conclusion. It would be useful to briefly describe again in what they differ.

AC S9. The "four formulations" of the retrieval are now mentioned under "Summary and Conclusions" as follows:

"Four formulations of the retrieval were used, based on either the SPARTICUS or TC4 field campaigns, with the smallest size-bin of the 2D-S probe either assumed valid ($N(D)_1$ unmodified) or by assuming $N(D)_1=0$. The SPARTICUS unmodified $N(D)_1$ assumption gives the highest N values while the TC4 $N(D)_1 = 0$ assumption yielded the lowest N values. The N predicted from these two formulations differed by a factor of two (see Fig. 5), thus defining a possible systematic uncertainty in N ."

Technical corrections:

1. p. 2 l. 13 and 16: it would be better to explicitly refer to "ice clouds" instead of "clouds"

Done

2. p. 11 l. 17: "limit is 1.031"

Done

3. p. 43 l. 15: "m.s⁻¹ in"

Text removed