

Atmos. Chem. Phys. Discuss., referee comment RC1
<https://doi.org/10.5194/acp-2022-401-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Review of "14 years of lidar measurements of Polar Stratospheric Clouds at the French Antarctic Station Dumont d'Urville" by F. Tencé et al., proposed for publication in Atmospheric Chemistry and Physics

Anonymous Referee #1

Referee comment on "14 years of lidar measurements of polar stratospheric clouds at the French Antarctic station Dumont d'Urville" by Florent Tencé et al., Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2022-401-RC1>, 2022

In this article, the authors apply several classification schemes on lidar measurements of polar stratospheric clouds above Dumont d'Urville. They compare the results of each scheme, as global statistics and as a function of temperature and altitude. They document how the choice of thresholds affect the retrieved PSC types (using a single classification scheme) on very fine scales during a case study observed over DDU. Finally, they attempt to derive trends of PSC occurrences above DDU over a 13-years period using ground-based lidar observations.

Investigating how the classification schemes impact the retrieved PSC types is an important and difficult exercise. Comparing the results of each classification scheme is interesting and useful, it highlights very well how the selections of threshold affect the relative importance of each PSC family and at what altitudes and temperatures. The case study makes a good job of showing how even very small changes in retrieved optical properties can switch a PSC between categories. Results from the trend study are interesting in light of the reported stratospheric warming in the southern hemisphere.

I have no major objection against the publication of this article. I have many small comments that I'd like the authors to address before publication.

Minor comments

General comment 1: The article is missing a good documentation of the ground-based lidar dataset it is built upon. What is the period of observation covered? How frequent, how long are observation periods? Are specific months (JJAS?) selected, and the rest

ignored? Are there annual/seasonal/hourly changes in operation and sampling? How do the sampling coverage and statistics compare to those of CALIOP? This alone could explain differences in ground-based vs spaceborne retrievals.

General comment 2: The text in most figures is small enough to be sometimes illegible. It would be better if most figures were displayed full-width, but the text would probably remain too small. This is particularly true for figures 1, 2, 3, 4, 7 and 8, in which number axes are extremely small. See the image in the supplement that shows some article text (up left) and a bit of figure 1. Please fix this and make text readable in figures.

- l. 17 (and others): here you refer to PSCs as "stacks of layers". My impression is that it is a very lidar-centric view. PSCs are 3-dimensional structures, as such they can be viewed as stacks of layers, but they could also be viewed as columns of vertical slices, arrays of cubes, etc. I'm not sure what this particular way of describing 3-dimensional structures brings to the table. Please clarify: is there something in the nature of PSC formation and dynamics that leads to a structuration of overlapping, horizontally-consistent slabs? (note this is definitely not true for wave PSCs)
- l. 46: Later... (Larsen, 2000). Please check the chronology of your paragraph here
- l. 63: "different set"
- l. 65: "we decided to consider 3 different classifications proposed by Blum, Pitts and an updated version of P11 is also considered" -- please fix phrasing: the updated version of P11 either is one of the three different classifications, OR is also considered, but not both.
- l. 65: "Following their conclusions": whose conclusions? Achtert and Tesche 2014 are quite far away, please clarify.
- l. 68-79: here only sections 1 to 3 are mentioned. Please include all sections. The lidar instrument is actually presented in Section 2, not section 1 as the text says. Processing and schemes are described in section 3 (not 2), etc.
- l. 83: "(NDACC"
- l. 96 and elsewhere: it looks like you've chosen to use "scattering" where I would have expected "backscattering". Is there a reason for this? Could you clarify in the text that this is your meaning?
- l. 96: you defined the backscatter/scattering ratio profiles as the ratio of total scattering to molecular scattering. Is any of those two attenuated? Please be explicit.
- l. 101: "data is" plural
- l. 110: "Each instruments"
- l. 137: "NCEP reanalysis product is the result of a cooperation between NCEP and NCAR": this info is already provided on lines 127-128.
- l. 151: I don't think β_{tot_perp} has been defined here yet. I'm guessing that each of the three groups [R_T , $R_{//}$] etc is used by a different classification scheme. Please make that explicit.
- l. 165: thanks for the very interesting reference to Behrendt and Nakamura, 2002. I could not find the 0.443% in the text of the article itself, could you expand a bit on how you obtained it? i.e. what temperature or other input parameters you've selected?
- l. 169: this was already stated line 151
- l. 171: "PSC classification is challenging as described in the introduction but critical": weird phrasing. Please rephrase as e.g., "As described in the introduction, PSC classification is challenging. It is, however, critical..."
- l. 176: "Achtert and Tesche..." the same sentence is already more or less present on page 3

- l. 199: MX1, MX2
- l. 211: It is unclear to me why you consider P11 in addition to P18. Isn't P18 supposed to supersede the P11 algorithm? Are there reasons why anyone who would like today to study PSCs using CALIOP measurements should go for the P11 algorithm? Version 2 of the CALIPSO PSC product is totally based on P18, so anyone who would like to study PSCs using CALIOP measurements is stuck with P18 anyway (unless she's willing to process the classification herself). Could you clarify what is the point of including P11 in the comparison?
- l. 237: "features"
- l. 242-244: unclear, what are you planning to do with those mixed-phased clouds? Are you going to make them appear as a separate entity, or subsume them in the category of the dominant particle type, or something else?
- l. 248: "A distribution of PSC types... published in Tesche et al was included": How did you get the numbers from Tesche et al. 2021? As far as I can tell, the article itself did not include numerical values for its retrievals, so did you lift numbers from the figures? If so, it is surprising you can reach precisions like 15.8%.
- l. 270-274: From what you write here I understand that ice PSC are under-represented in DDU lidar PSC observations. If that is indeed what you meant, could you please spell it out explicitly? This actually could be checked (relatively) easily -- in each CALIOP profile one could see for a given PSC type the frequency of opaque tropospheric clouds underneath. According to your explanation, opaque tropospheric clouds should be relatively more frequent in presence of ice PSC than in presence of other PSC types. If you think this is outside the scope of the present paper, perhaps mention it as a possible perspective.
- l. 272: "Marginal" According to your discussion, CALIOP results should be closer to the correct number of ice PSC, and they report a frequency of 16% for ice PSC. Is that marginal?
- l. 301: It would be interesting to apply the various classification schemes on the entirety of the CALIOP observations, and identify in what geographical regions the results diverge. This is clearly outside the scope of the current paper ☐☐☐
- The discussion of the comparison suggests to me that outputs of classification should come with some kind of reliability indicator, that would decrease as the measured optical parameters get closer to category boundaries. Such an indicator would improve comparisons and make inconsistencies between retrievals perhaps less significant. Is something similar already present in any product? If you think this is a good idea, you could take the opportunity to suggest it in your paper.
- l. 327: "This high variability must be kept in mind": why?
- l. 328: "horizontal smoothing... due to the transport" the transport of what? Please clarify.
- l. 333: the type changes throughout the whole day, not just once at 5PM. But your point stands.
- l. 334: Related to my previous point about a type reliability indicator, do the optical parameters of this cloud hover near the boundary between two categories in the classification diagram? Would an indicator help identify this situation and flag it as unreliable?
- l. 339-341: Could you specify if, in your opinion, these changes in composition (derived from the changes in optical properties) are consistent with the speed of the deposition and growth processes that would drive the change in composition? In other words, are the changes in composition trustworthy, or are they a demonstration of the limitations of the optical classification approach?
- l.354: Here by "lidar" you imply an HSR-capable lidar. Please clarify.
- Figure 5: Here the labels are quite readable, but the decision to make the figure wide and short makes it very hard to identify any structure visually (especially in Figure 5a). Could you please reorganize the figure to change its aspect ratio somehow? Maybe make it a 3-columns/1-row full-width figure?
- l. 366-367: "To investigate the effect of temperature variation on PSC..." do you mean

- "the impact of the choice of temperature dataset on the results of PSC classification"?
- Figure 8: I'm sorry but I don't understand what is being shown here. As I understand it, the figure shows three numbers : A) the number of days in which the lidar observed a PSC (red triangles), B) the number of days in which the ERA5/NCEP temperature allowed PSC formation (green/red crosses), and C) the number of days in which ERA5/NCEP temperatures were 2K below the TNAT formation threshold, AND no lidar measurements were available (grey arrows). In my view, "the number of days in which the ERA5/NCEP temperature allowed PSC formation" is the same as "the number of days in which ERA5/NCEP temperature were 2K below the TNAT formation threshold". In that case, A+C should be equal to B. This is clearly not the case in the figure, so I must have misunderstood something, but I can't find elements in the text to clarify my misunderstanding. Please help.
 - l. 408: the negative trend that is found here mostly depends on the reliability of ERA5 and NCEP stratospheric temperatures, and on the presence of a overall stratospheric temperature trend in those datasets, correct? Could you make it clearer why your results are not just confirming the presence of a warming trend in ERA5/NCEP stratospheric temperatures? i.e. what is the lidar bringing here?
 - l. 445-448: I understand from your conclusions that 1) applying the three classifications schemes to ground-based lidar observations leads to results that agree quite well, and 2) applying the same classification scheme (P18) to ground-based and spaceborne lidar leads to results that agree well too. From this, I understand that the choice of classification scheme has after all little importance on the results. Do you share that opinion? If not, could you amend your conclusions to include arguments for the opposite viewpoint?

Please also note the supplement to this comment:

<https://acp.copernicus.org/preprints/acp-2022-401/acp-2022-401-RC1-supplement.zip>