Review of acp-2022-40: **14 years of lidar measurements of Polar Stratospheric Clouds at the French Antarctic Station Dumont d'Urville,** by Tencé et al.

This paper presents analyses of a 14-year database of PSC observations from the ground-based lidar at the French Antarctic station Dumont d'Urville (DDU). After describing the lidar system at DDU, the authors present an extensive evaluation of three published PSC classification schemes they consider for use with the DDU database. One of the classification schemes is then selected to use to illustrate a sample PSC event of interest. Next, the authors evaluate the viability of several temperature datasets (radiosondes, ERA5, NCEP, and IASI) for use in analysis of the DDU PSC data. Their analyses indicate that the ERA5 temperature data is best suited for use in processing and interpretation of the DDU PSC data. Finally, the authors derive a PSC trend estimation utilizing temperature statistics as a surrogate for PSC data due to sampling issues in the ground-based lidar database. Their results show a decreasing trend of -5.7 PSC days per decade based on 14 years of data (2007-2020).

The overall goals of the paper are not clearly articulated in the abstract or introduction and remain unclear to this reviewer. A significant portion (almost 50%) of the paper is dedicated to the evaluation of the three published PSC composition classification schemes, B05 (Blum et al., 2005), P11 (Pitts et al., 2011) and P18 (Pitts et al., 2018). Although the authors state that the purpose of their study is "not to review, rank or assess the relevancy of these classifications … but rather to find a proper framework to analyse our DDU lidar measurements," it seems that reviewing and assessing the classifications is indeed a primary focus of the paper. After the extensive discussion of the three classification schemes, the authors conclude that the three are comparable, but select the P18 scheme to examine a sample PSC event of interest. If a detailed comparison of the schemes is not a focus of the paper (and I don't think it should be), then it would be more succinct to simply describe the classification approach the authors think most appropriate and then present more detailed analyses of the 14-year PSC database using this classification. What new insights into PSCs can be gleaned from the 14-year DDU database? For instance, instead of showing just one sample PSC event of interest, the authors could present some multi-year statistics on the mesoscale characteristics of PSCs. This could be a unique contribution from the DDU dataset.

The discussion of the temperature datasets could also be much more succinct providing the justification for why the ERA5 data is chosen to be used in the DDU data processing and analyses. Again- depends on what are the primary goals of the paper.

The trend analysis is interesting, but the inclusion of the CALIOP data here (see detailed comments below) could provide additional insight.

As I'm sure the authors are aware, there are at least two other recent published studies of PSCs using ground-based lidars located in the Antarctic:

Snels et al. (2021). Quasi-coincident observations of polar stratospheric clouds by ground-based lidar and CALIOP at Concordia (Dome C, Antarctica) from 2014 to 2018, Atmospheric Chemistry and Physics, 21, 2165–2178.

Snels et al. (2019). Comparison of Antarctic polar stratospheric cloud observations by ground-based and space-borne lidar and relevance for chemistry-climate models. Atmospheric Chemistry and Physics, 19, 955–972.

How do the DDU PSC statistics compare with these other datasets? I would expect the papers to at least be cited somewhere in this manuscript.

Then finally, there are numerous instances of awkward and/or confusing text throughout the manuscript. I have pointed out many of these below and suggest some wording changes. Although individually they are generally not major issues, the large number of these instances make the manuscript difficult to follow and should be revised where appropriate.

In conclusion, although I do have issues with many portions of the manuscript, I believe with some significant revisions and possible inclusion of new analyses (as described below), it may be suitable for publication in ACP.

Specific Comments:

1) The current Introduction in Section 1 needs major reworking. What are the main goals of this study? How are the ancillary datasets used? These need to be clearly articulated here. This will provide a roadmap for the rest of the paper. From the current introduction, it's not clear why so much effort is going into evaluating the classification schemes. Is this a major goal of the paper? The last paragraph of the current Introduction is attempting to describe the remaining sections of the paper- but it is not accurate or complete. Please rewrite to better summarize what is in each subsequent section (e.g., 2.-Methods, 3- Lidar data processing, 4- Results, 5-Conclusions).
2) Line 3: The meaning of the term "tight model parameterization" is not clear. Do you mean mathematically simple and/or computationally fast?
3) Line 17: What is meant by "stacks of layers featuring different mixtures." Please cite a reference and describe in a little more detail what is meant by this phrase.
4) Line 18: change "when temperature" to "when the temperature"
5) Lines 20-22: Strictly speaking, I believe denitrification and dehydration refer to the redistribution and **irreversible** removal of HNO3 and H2O from the stratosphere. Uptake of HNO3 and H2O by itself (through particle formation) may be reversible. Therefore, denitrification and dehydration occur by sedimentation of large NAT or ice PSC particles that contain HNO3 and/or H2O.
6) Line 23: The phrase "a lot" is not a good choice for a technical paper. Would be good to have some citations here on what significant improvements have taken place and what remains to be understood. Perhaps more relevant here, is this study going to improve our understanding of any of these outstanding questions?

7) Line 27: “sulfur aerosols” should be “sulfuric acid aerosols”
8) Line 27: meteoritic material- is there a citation you could include here that shows meteoritic material may be efficient PSC nuclei?
9) Lines 27-29: I suggest listing relevant citations in the same order as the particle compositions (ice, NAT, STS) … (Peter and Grooß, 2012; Hanson and Mauersberger, 1988; Carslaw et al., 1997). Did Peter and Grooß (2012) actually perform lab studies on ice? I thought this was a chapter in a book.
10) Lines 30-31: “NAT particles only nucleate on pre-existing particles.” - what pre-existing particles? ice? meteoritic material? Citations?
11) Line 37: The Wegner et al. (2012) Figure 1 only shows efficiencies for liquid aerosol (binary and ternary) and NAT, not ice. Aren't these efficiencies primarily based on the available surface area? Is it really composition dependent or mostly surface area density dependent?
12) Lines 38-39: This sentence is confusing to me. What do you mean by “pure STS, NAT, and ICE blends of chemical compounds?” Are you simply referring to the chemical makeup of the particles?
13) Line 40: “Poole and McCormick (1988) in 1988.” I think it is obvious that Poole and McCormick was published in 1988, so you don’t need the additional “in 1988”
14) Line 41: “set” should be plural “sets”
15) Line 43: “Achtert and Tesche (2014), 2018).” I think this is a typo- “, 2018)” should be deleted.
16) Line 45-46: “… but whose presence was not proven in atmospheric observations.” Suggest rephrasing as “but has yet to be confirmed by atmospheric observations.”
17) Lines 46-48: What studies have shown these “stacks of fine layers?” I don't think that the Larsen paper shows that PSCs often occur as stacks of layers of different particle types- at least I didn’t see any mention of this in the report. Larsen does conclude that the temperature history of the air mass must be known to properly simulate the particle formation.
18) Lines 48-49: Sentence is poorly worded. Suggest something like “In addition, the temperature cooling rate is an important variable driving orographic PSC formation in both the Arctic and Antarctic (Noel and Pitts, 2012).”
19) Line 51: “only few” should be “only a few”
20) Lines 51-54: This sentence is not clear and too long. What is based on “optical properties”? The complex observational patterns? Surely not the parameterization schemes? Numerous phrases that are not clear: “tight as possible”? “observations derived patterns”? Please try to reword.
21) Lines 54-56: This sentence seems repetitive with the sentence L.40-42. Maybe you can combine this with the sentence on L. 40-42 and list the citations there?
22) Line 63: “different set of” should be “different sets of”
23) Lines 65-67: Years inside the parentheses are not necessary. Suggest rewording “… Blum et al. (2005) (hereafter called B05), Pitts et al. (2011) (hereafter called P11), and Pitts et al. (2018) (an updated version of P11, hereafter called P18).”

24) Consideration of P11: The P18 algorithm corrected several know deficiencies in the P11 algorithm. I believe the P18 has replaced P11 as the operational algorithm used to produce the CALIOP v2 PSC data products. Therefore, there is no reason to include the P11 version in your evaluation unless you just want to compare the differences between P11 and P18 (that was done by Pitts et al., 2018). Is that your goal here?
25) Line 70: suggest changing "station hosts" to "station has hosted"
26) Summary paragraph beginning on Line 68: As mentioned above, this last paragraph in the Introduction needs completely rewritten. This paragraph is attempting to describe the following sections of the paper- but is not accurate or complete. Please rewrite to better summarize what is in each subsequent section.
27) Line 80: Section 2 Methods: This section really doesn't describe methods- rather just the datasets used in the study. Probably should rename "Datasets"?
28) Line 82: Suggest changing "Since April 1989, an aerosol/cloud lidar system is in operation at DDU …" to "An aerosol/cloud lidar system has been in operation at DDU since April 1989 …"
29) Line 83: Add closing parenthesis after NDACC
30) Line 84: Delete extra space after "Antarctic atmosphere"
31) Lines 85-87: Awkward grammar- suggest rewording this sentence to "Although the measurement calendar focuses on the PSC season with nighttime setup, the recent focus on aerosol plumes either originating from volcanic or biomass burning activity (Tencé et al., 2022) has extended the measurement calendar to the summertime."
32) Section 2.1 DDU Lidar description: You say that the lidar capabilities have been continuously upgraded and cite the David et al. (2012) paper. Have there been any notable upgrades in the past 10 years since the David et al. paper?
33) Section 2.1 DDU Lidar description: Since you are introducing most if not all of the lidar optical parameters here, I suggest you move the equations defining the lidar parameters in Section 3.1 to this section. It seems more appropriate to have the definitions here. Maybe after L. 94?
34) Line 97: "saturation effects" - Would you please describe what the saturation effects are and add more detail on how they are removed?
35) Line 98: "homogeneity of the scene" –How is the homogeneity quantified and used?
36) Line 100: "altitude" should be plural "altitudes"
37) Section 2.1.1 IASI temperature product: What is the main role of the IASI temperature product in this study? Should note that in the introduction (section 1).
38) Line 111: "instruments" should be singular "instrument"
39) Line 112: à should be "a"
40) Line 113: delete period after PM
41) Line 118: "temperatures" should be "temperature"
42) Line 120: "very good agreement" – Please be more quantitative- what is very good agreement?
43) Lines 124-125: Sentence is worded awkwardly. Suggest rewording something like "As discussed in further detail in the following sections, reanalysis temperature products are

often utilized to complement or replace local radiosonde measurements for both data processing and interpretation of ground-based lidar measurements."

44) Lines 129-130: What do these acronyms (4D-Var, Cy41r2) mean?
45) Line 133: "interpolated at DDU location" – "interpolated to the DDU location" Is it simply linearly interpolated from the original product grid (0.25 x 0.25 degree)?
46) Line 135: "dynamic tropopause"- how do you define the dynamic tropopause? Is this an ERA5 product that you interpolate to the DDU location?
47) Line 139: NCEP product: What is meant by "provide an output for DDU"?
48) Reanlyses Data discussion in general: Again, there should be some discussion in the Introduction of how the reanalyses data will be used in this study. What are the uncertainties in reanalyses data products?
49) Section 3.1 PSC detection by lidar: I found the description of the PSC detection here to be confusing. The first step is some "pre-processing" that identifies time segments that contain aerosol/cloud? What is dynamic time averaging? What do you mean by "next step summation according to homogeneity"? What is the "peak detection algorithm"? Are you just searching each profile identified as containing aerosol/cloud for peaks that identify layers? The output of the detection algorithm are profiles of lidar parameters with one or more layers identified as being aerosol/cloud? You mention in Lines 150-152 that a type is attributed to each layer. Isn't the composition classification performed separately from detection and dependent on the specific scheme being used as described below in Section 3.2? Please try to rewrite this section more clearly with more detail.
50) Line 144: "assuming no particular extinction" – do you mean "assuming no particulate extinction"?
51) Lines 155-169: As mentioned above, I suggest you move these lidar parameter definitions up to Section 2.1
52) Section 3.2 Classification schemes: General comment- I think too much emphasis has been put on discussion and evaluation of the classification schemes. You state that the purpose of the study is not to "review, rank or assess" the classification schemes, but to "find a proper framework" to analyse your data. It sure seems that you are reviewing and assessing the classifications. What does "proper framework" mean? How do you decide which classification scheme provides the proper framework.
53) Lines 176-182: Much of this was already discussed in the Introduction.
54) Line 225: "10^2 order or magnitude" should be "10^2 order of magnitude" Do you really mean 100 orders of magnitude or just 2 orders of magnitude (factor of 100 in magnitude)?
55) Lines 230-232: I agree that it doesn't make sense to use the MLS measurements directly. But wouldn't be better to use a climatology of HNO3 and H2O and have a time dependent threshold? It would be straight forward to produce a climatology from the MLS data.
56) Line 235: Suggest rewording this to read: "The wave ice category defined in P11 and P18 was ignored in this study ..."
57) Line 236: verb tense doesn't match- "… published classifications … features" should be "…published classifications … feature"

58) Line 242: Suggest changing "disequilibrated" to "non-equilibrium"
59) Discussion of Figure 2: Figure 2d (derived from Tesche et al., 2021) is based on only two Antarctic seasons (2012 and 2015) and I believe uses only a subset of CALIOP measurements randomly selected to represent the possible sampling of a ground-based lidar that is affected by cloudiness and other measurement-inhibiting factors. What is the relevance of this figure to the others in Figure 2 that are based on 14 years of data? Doesn't seem to be a fair comparison. It certainly would be straight forward to derive a new figure using the CALIOP data for the same timeframe as your DDU data- then the comparison would be more meaningful.
60) Discussion of ice discrepancy in Lines 271-274: To better investigate this- I suggest you subset the CALIOP data to the DDU location and evaluate the ICE abundance on days when the DDU lidar operated versus days in which DDU lidar didn't operate.
61) Lines 280-283: These two sentences are not clear- not sure what you're trying to say.
62) Lines 285-286: Why would optical properties be dependent on latitude?
63) Line 290: "barely never" – that phrase makes no sense. Do you mean "barely ever"? Probably would be better to just say "rarely detected."
64) Figure 4 and corresponding discussion: Using threshold temperature values calculated with fixed values of HNO3 (10 ppbv) and H2O (5 ppmv) will likely produce misleading results. These values may be appropriate for early season (at ~50 hPa), but clearly are not representative for the bulk of the season after denitrification and dehydration have occurred. In reality, the gas phase abundances are much lower over most of the season and the threshold temperatures will correspondingly be lower. It would not be too difficult to derive a climatology of HNO3 and H2O from MLS data that reflects the seasonal and altitude variation and then use this to calculate time dependent temperature thresholds. This would provide a much more realistic evaluation of the PSC detections versus altitude and temperature. But then I ask, is the analysis presented in Figure 4 even necessary for this paper? What is the purpose of this analysis?
65) Section 4.2: General comment- why only show one sample PSC event? Now that you have selected a classification scheme- why not process all 14 years of data and produce statistics on interesting aspects of the PSCs such as the mesoscale characteristics? One example is OK- but how representative is it? A statistical analysis would be very interesting and much more compelling for inclusion in the paper. Can you do this?
66) Line 320: Figure A2- I think you are actually referring to Figure A1 here. Why put the model analysis in an appendix? I think it is OK to include in the main text.
67) Line 324: Again think you mean Figure A1
68) Line 325: "fully validate the model" is a strong claim- maybe "The model produces PSC at the 435 K and 475 K levels, and no PSC at the 550 K level above DDU, consistent with the lidar measurements."
69) Line 332: "… temperature history cannot be accounted for." I don't think there is a limitation that temperature history cannot be used and in fact future schemes may indeed include temperature history as a parameter. Therefore, I suggest rewording as "… temperature history has not been accounted for."
70) Line 334: RT and RICE should be written as R_T and R_{ICE}

71) Figure 5 discussion: Are the Tice thresholds based on 5 ppmv H2O? This is not likely representative of 28 August when the stratosphere may be severely dehydrated. How would this change your interpretation?
72) Figure 5: hard to see the temperature differentials in the bottom panel- would be helpful to reduce the range of the color bar.
73) Figure 7 discussion: I assume the temperature differences are based on the reanalyses data from over DDU (interpolated) and the temperature at the true location of the radiosonde downwind from DDU- right?
74) Line 379: "NCEP is obviously less accurate …" Is it obvious? Differences with the radiosonde are larger- but doesn't necessarily imply NCEP is wrong. Do you have a citation that NCEP is less accurate than ERA5 and IASI?
75) Lines 388: This statement should appear in the Introduction and help define the focus of the paper!
76) Line 395: "Both NCEP and ERA5 …" You concluded in the previous section that the NCEP temperatures are not as accurate as the ERA5 and therefore the ERA5 data would be used in this study. Why include the NCEP data here?
77) Lines 399-406: How was T_{NAT} calculated for the trend analyses? Did you use a fixed value of 10 ppbv over all altitudes and days? Would a more accurate value reflecting denitrification change your results?
78) Trend analyses: Did you consider subsetting the CALIOP PSC data from 2007-2020 to the DDU location and compare this with the temperature time series? This would be a very interesting exercise to include here.
79) Line 435: "… ERA5 slightly overestimates …" This conclusion is based on comparisons with radiosondes not necessarily collocated with DDU due to balloon drift. Doesn't the drift of the balloon make it difficult to conclude anything quantitative about accuracy since there are spatial temperature gradients that introduce differences?
80) Figure 9: I find the figure confusing- but maybe I just don't understand what is being shown. For PSC thickness between about 2-7 km, the $T<T_{NAT}$ domain is smaller than the PSC thickness? But you state that the figure shows that the $T<T_{NAT}$ domain is significantly larger than the actual PSC thickness! What am I missing?
81) Section 5 Conclusions: No specific issues at this point. Will reserve comment for the anticipated revised version.