We thank the reviewer for their comments. Please find below our point by point response in blue.

We draw the reviewer's attention to the fact that, following a similar comment of all reviewers, a figure 1 has been added to the manuscript, featuring information on the operation statistics of DDU lidar. Therefore, all figure numbers have been incremented accordingly as compared to the first version of the manuscrit.

Also, to address another comment, the threshold temperatures T_{NAT}, T_{STS} and T_{ICE} are now computed based on the closest MLS H_2O and HNO_3 concentration measurement. Although this does not change their meaning or their interpretation, most of the figures have gone through some slight changes.

Concerning the comparison of our PSC record with other Antarctic groundbased datasets, it is right that it should have been included and it is now the case. DDU PSC distribution is compared with PSC types distributions produced at McMurdo and Concordia by Snels et al., (2019, 2021), a new paragraph has been added in the manuscript, lines **341-354**:
"At McMurdo antarctic station (77.85°S - 166.66°E), from 2006 to 2010, Snels et al. (2019) reported a mean distribution of 13.8% STS, 71.6% NATmix, 2.6% ENH and 12% ICE. During this period, CALIOP observed approximately 10% more STS and 10% less NATmix and otherwise shows a good overall agreement with the groundbased lidar. Snels et al. (2021) compare Concordia groundbased PSC detections to CALIOP measurements around the station from 2014 to 2018. This study mainly focuses on the agreement between groundbased and spaceborne instruments and do not directly provides the PSC types distribution observed at Concordia by the local lidar, but still shows the distribution observed by CALIOP around the station. From 2014 to 2018, the yearly occurrences of STS represent from 14 to 38%, NATmix from 42 to 67%, ENH from 5 to 11% and ICE from 10 to 28%. The distribution shows a high annual variability, but we still can point out differences with DDU PSC types distribution. Both McMurdo and Concordia measurements feature a higher proportion of ICE detections and less STS observations as compared to DDU. The 38% STS share observed at Concordia in 2014 is considered to be an outlier. This is consistent with the results of Fig. 19 of Pitts et al. (2018) which show that main area of ICE occurrence is located inside the continent while DDU is located on the Antarctic coast. The temperature necessary to form ICE crystals are reached less often at DDU, on the edge of the vortex."

Following several comments, a comparison with CALIOP PSC measurements around DDU is now included in the study and is extensively described in this document. It notably relies on the method used by Snels et al. (2021) for Concordia and this study is therefore cited in this context.

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).
The introduction has been edited to expose more clearly the goals of our study, lines **80-85**:
"Considering the latest laser source replacement in 2005 and the continuous monitoring from 2006, DDU PSC dataset is compared to the spaceborne PSC measurements conducted by CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) in the vicinity of the station, from 2007 to 2020. Two of the major roles of a ground station are to perform process studies and establish decadal trends. Such trends are highly valuable because they reflect the evolution of the stratosphere, in terms of temperature and chemical compositions. In this study, DDU lidar measurements are exploited to produce a trend of PSC days per year at DDU."

The last paragraph has been rephrased to feature more clearly the description of the remaining sections. Edited, lines **85-93**: "Section 2 presents the data and instruments exploited in this study. Section 3 introduces the PSC detection methods as well as the considered classification schemes. Section 4 exposes the results of the study. First, the outcomes of the application of the classification schemes B05, P11 and P18 to the DDU lidar data record are presented and discussed, and CALIOP and DDU PSC measurements are compared. Then the analysis of an interesting example of a long lidar session is used to illustrate the unique capabilities of lidar measurements in characterizing very fine vertical features in PSC fields. Third, lidar measurements and temperatures from ERA5, NCEP and Infrared Atmospheric Sounding Interferometers (IASI) are combined to produce a PSC occurrence trend from 2007 to 2020. To support the use of these temperature data sources, they are compared to temperatures from radiosondes launched daily at DDU. Finally, some challenges of PSC parametrization in climate models are discussed before exposing the main conclusions."

2) Line 3: The meaning of the term "tight model parameterization" is not clear. Do you mean mathematically simple and/or computationally fast?
"Tight model parametrization" is meant here in the sense that models need to keep the number of variables used for PSC parametrization as low as low possible.
Tight actually refers to a low number of model variables for parametrization of PSC microphysics. It is following the reviewer's guess related to the degree of mathematical/physical complexity. For clarity edited in the manuscript, lines **3-4** : "...and model parametrization constraints".

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.

We mean that PSC fields often consist of several layers corresponding to different types of clouds, i.e. different mixtures, thus, different chemical composition. Considering comments from below, layer is defined as consecutive vertical bins of the same type. Illustration is clear on Figure 7 in our paper, Figure 4 in Blum et al.[2005] or Figures 4 to 8 in Snels et al. [2021].
Since this point was raised by several reviewers and was not central to our analysis, it was rephrased as follows, line **18**: "often observed as layers featuring different chemical compositions"

4)Line 18: change "when temperature" to "when the temperature"
Edited: "when the temperature" line **19**

5) Lines 20-22: Strictly speaking, I believe denitrification and dehydration refer to the redistribution and irreversible removal of HNO_3 and H_2O from the stratosphere. Uptake of HNO_3 and H_2O by itself (through particle formation) may be reversible. Therefore, denitrification and dehydration occur by sedimentation of large NAT or ice PSC particles that contain HNO_3 an/or H_2O.
That is what we meant but the sentence was poorly phrased, edited the manuscript to be clearer. What was meant is that HNO3 / H2O are taken from the gas phase into PSC particles, which are then removed from the stratosphere as the PSC sediment.
Edited line **22**: "Denitrification and dehydration, mostly through the uptake of HNO3 and H2O by PSCs and the subsequent PSC sedimentation, decrease HNO3 and H2O stratospheric concentration and hence enhance ozone depletion."

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?

The paragraph was edited to include examples of the improvement of our global understanding of PSCs in recent years as well as examples of the PSC mechanisms still actively studied.
Edited lines **23-32**:

"Despite major improvements in the recent years due to enhanced research and monitoring capabilities, some PSC-related aspects are still to be understood. Spatial measurements brought a global point of view able to grasp the spatial and temporal distribution of PSC during winter (Pitts et al., 2018). Studies highlighted the need to take wave-induced temperature variations in account to adequately model PSC occurrences (Cairo et al., 2004; Höpfner et al., 2006; Eckermann et al., 2006; Engel et al., 2013; Tritscher et al., 2019). However, some PSC particles formation pathways are still debated (Tritscher et al., 2021) and the adequate model parametrization of PSC is still challenging. Besides, the recent stratospheric injections of aerosols caused by volcanic eruptions and wildfires also raise questions on the potential interaction with PSC formation processes and subsequent stratospheric ozone depletion (Tencé et al., 2022; Ansmann et al., 2022; Rieger et al., 2021 Stone et al., 2021)"

7) Line 27: "sulfur aerosols" should be "sulfuric acid aerosols"
Edited: "sulfuric acid aerosols" line **35**

8) Line 27: meteoritic material- is there a citation you could include here that shows meteoritic material may be efficient PSC nuclei?
This is still a subject of discussion, so we moderated our assumption. Section 3.3 in Tritscher et al. (2021) discusses the possibility of meteoritic material to act as nuclei for PSC particles and the mutliple laboratory and fields works published on that topic. We can for example cite the work of James et al. (2018) who proved that meteoritic material can trigger NAT nucleation in PSCs, or the field measurements of Ebert et al. (2016) which suggest meteoritic material could be involved in a PSC nucleation pathway.
Edited lines **35-36**: "or to a lesser extent meteoritic material, which role in PSC particles nucleation is still a subject of discussion (Ebert et al., 2016; James et al., 2018)".

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.
Peter and Grooß (2012) is indeed a book chapter notably reviewing PSC particles nucleation pathways. Koop et al. (2000) may probably be more suited.
Edited line **38**: "'(Koop et al., 2000; Hanson and Mauersberger, 1988; Carslaw et al., 1997)"

10) Lines 30-31: "NAT particles only nucleate on pre-existing particles." - what pre-existing particles? ice? meteoritic material? Citations?
Original intent was to highlight that NAT only nucleates heterogeneously because the homogeneous nucleation is kinetically suppressed (Koop et al., 1995; Koop, Carslaw et al., 1997), and not to discuss the potential nuclei here. NAT nucleation on pre-existing ice particles is established (for example, Koop, Carslaw et al., 1997) but does not explain all NAT observations (Drdla et al., 2002; Pitts et al., 2011).
Citations added line **40**: (Koop et al., 1995; Koop, Carslaw et al., 1997, James et al., 2018)

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?

Ice particles are indeed missing from this figure. I mistook it with the adaptation made by Tritscher et al. (2021) that includes ice. I changed the citation.
Edited line **46**: "in Fig. 39 of Tritscher et al. (2021)".

Efficiencies are indeed surface area dependent, the point of Fig. 39 of Tritscher et al. (2021) and Fig. 1 Wegner et al. (2012) is to link particle types, and therefore composition, to chlorine activation efficiencies. I edited the previous sentence to mention the link with surface area density.

Edited line **42**: "Depending on their dominant type of particles, different PSCs have different surface area densities and chemical heterogeneous reactivities..."

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?
Original sentence is indeed misleading. Intent was to say that PSC classifications were based on the optical properties of pure STS, NAT and ICE along with mixtures of the latter species. Edited lines **47-49**: "From these three basic particle types and their combinations, more detailed types of PSCs were identified for the purpose of creating classifications based on optical properties of the pure STS, NAT and ICE blends of chemical compounds."

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"
This is of course a typo in the bibliography management, it has been addressed throughout the revised manuscript.

14) Line 41: "set" should be plural "sets"
Edited "sets" lines **50**

15) Line 43: "Achtert and Tesche (2014), 2018)." I think this is a typo- ", 2018)" should be deleted.
It is indeed a typo, it was deleted.
Edited: ", 2018)" was deleted line **53**

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."
Rephrased as suggested.
Edited: "but have yet to be confirmed by atmospheric observations." Lines **56**

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.
The Larsen paper is cited as support of the statement on the history of the air masses.
See comment 3) concerning the stacks of fine layers statement. Since the mention of the stack of layers was not useful here, it is removed to avoid confusion.
Edited line **56**: "Studies showed that PSC type identification depends on the history of the air masses due to hysteresis effects in PSC formation along the temperature scale (Larsen, 2000)."

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)."

Rephrased as suggested, lines **58-59**.
"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"
Corrected. Edited "only a few" line **61**

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.
Sentence actually needs rephrasing. Edited at lines **61-63** into: "Therefore, the combination of model constraints and high rate of unclassified observations (due to either instrumental concerns or unequilibrated particles) led to some redefinition of the boundaries between the existing PSC classes rather than considering additional classes."

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?
The citations were moved in the previous paragraph, lines **50-51**, and the sentence lines **63-64** was rephrased. The idea was to mention that these classifications are based on lidar measurements, whereas the previous paragraph focused on the evolution of PSC classifications following the evolution of our understanding of these clouds.

Edited: "The PSC classifications schemes previously mentioned are based on ground-based or space-borne lidar measurements" Lines **63-64**

22) Line 63: "different set of" should be "different sets of"
Corrected. Edited "different sets of" line **72**

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)."
Typo with the bibliography corrected. The years in the parentheses line **76** were removed.

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?
We consider P11 as it is involved in the analysis and conclusions of Achtert et al. (2014), stating B05 and P11 are comparable when applied to the Esrange lidar database with satisfying result. When applying P18 to our groundbased dataset (formerly using P11), we had to adapt the classification so comparing the outcomes of B05, P11 and P18 at DDU sounds consistent to us. It is also a way to validate the adjustment of P18 thresholds to our groundbased setup.
Moreover, as P11 and P18 as well as their intercomparison in Pitts et al. (2018) are based on CALIOP measurements, we consider the additional use of a groundbased dataset interesting. Finally, the optical properties used in P11 and P18 are different, so it is interesting to us to have both.

25) Line 70: suggest changing "station hosts" to "station has hosted"
Edited line **79** "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.

The paragraph is rephrased to better introduce the sections content. Please see comment 1 with the added manuscript discussion in lines **86-94**.

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"?

It is true, the section was renamed "Data and instruments", line **95**.

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 ..."

Rephrased as suggested, line **97**: "An aerosol/cloud lidar system has been in operation at DDU since April 1989 ..."

29) Line 83: Add closing parenthesis after NDACC

Edited ")" Line **98**

30) Line 84: Delete extra space after "Antarctic atmosphere"

Extra space deleted

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."

Rephrased as suggested, lines **100-102**: "Although the measurement calendar focuses on the PSC season with nighttime setup, recent work on aerosol plumes either originating from volcanic or biomass burning activity (Tencé et al., 2022) suggests extending 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?

Notable technical upgrades concern the redesign of the eclectronical synchronization circuit, software acquisition cards, photomultiplier tubes and emission optical bench. Noticeably, the laser source has been renewed last 2022 summer campaign, and redesign of the channel box is under study.

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?

It is indeed more adapted. Variables definitions were moved from section 3.1 to section 2.1. Consequently, two sentences were displaced and one was added lines **114-115** before the variables definitions: "The different lidar related variables used in this study are defined as follows:".

34) Line 97: "saturation effects" - Would you please describe what the saturation effects are and add more detail on how they are removed?

The mention of saturation effects misses context elements: basically, saturation only concerns a very small fraction of the 2007-2020 individual profiles and is mostly related to the new setup with

the laser replacement last year. Saturation leads to a bias in the overall shape of the lidar profile due to photons being missed at the maximum power altitude. Correcting saturation effect is only done with a clear tropopause: indeed, thick cirrus layers will prevent us from rebuilding the molecular shape of the lidar signal. Saturation correction parameters are set from expected Rayleigh scattering calculations and maximum saturation value of the photomultiplier tube.

35) Line 98: "homogeneity of the scene" –How is the homogeneity quantified and used?
The statement was too vague because these pre-processing steps are described in section 3.1, we provided more details. Please see comment 49) which extensively addresses this topic.

36) Line 100: "altitude" should be plural "altitudes"
Corrected. Edited "altitudes" line **132**

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).
The temperature retrieval from the IASI instrument is recent and provides a sample of daily temperature profiles extracted above the station. This feature is critical to us because it may be more suited to groundbased validation than other spaceborne datasets we could have used with lower relevancy in terms of distance, and measurement frequency. Besides, considering the input of another reviewer, IASI temperature dataset was included in the PSC trend presented in Figure 9. IASI is expected in this study to fill the observational role as temperature provider whereas NCEP and ERA5 are reanalyses. For example, IASI can detect small scale temperature variations relevant to PSC formation that reanalyses product may miss.

38) Line 111: "instruments" should be singular "instrument"
Corrected. Edited "instrument" line **151**

39) Line 112: à should be "a"
Corrected. Edited "a" line **152**

40) Line 113: delete period after PM
Corrected. Period deleted line **153**

41) Line 118: "temperatures" should be "temperature"
Corrected. Edited "temperature" line **158**

42) Line 120: "very good agreement" – Please be more quantitative- what is very good agreement?
Precisions have been added, from Bouillon et al. (2022). Edited, lines **161-164**: "Bouillon et al. (2022) showed that daily zonal mean differences between IASI ANN and ERA5 at mid and high latitudes are lower than 0.5 K between 750 and 7 hPa and reach 2 K at 2 hPa. Comparing IASI ANN with a global radiosoundings archive (Analyzed RadioSoundings Archive), Bouillon et al. (2022) found no significant bias and a standard deviation between 1 and 2 K for the Antarctic region."

43) Lines 124-125: Sentence is worded awkwardly. Suggest rewording something like "As discussed in further detail in the following sections, reanalysis temperature products areoften utilized to complement or replace local radiosonde measurements for both data processing and interpretation of ground-based lidar measurements."
Rephrased as suggested, lines **167-169**: "As discussed in further detail in the following sections, reanalysis temperature products are often used 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?

Cy41r2 is actually the usually refered to name of the Integrated Forecasting System (IFS) model cycle used by ECMWF.

4D-Var stands for "four-dimensional variational data assimilation". It is now specified in the manuscript. Edited line **173** "four-dimensional variational data assimilation"

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)?

It is indeed linearly interpolated from the original grid.

Edited line **177**: "to the"

46) Line 135: "dynamic tropopause"- how do you define the dynamic tropopause? Is this an ERA5 product that you interpolate to the DDU location?

The tropopause is defined as the lowest point between the 380 K potential temperature and |PV| = 2PVU. As the potential vorticity is not available in ERA5, it is calculated from the vorticity, the winds and the temperature.

This is now made clear as follows, lines **180-181**: "dynamic tropopause, defined as the lowest point between the 380 K potential temperature and |PV| = 2PVU (calculated from the vorticity, the winds and the temperature from ERA5)"

47) Line 139: NCEP product: What is meant by "provide an output for DDU"?

The NDACC/NOAA archive mentioned in the data availability section used for NCEP data provides daily profiles of temperatures and geopotential heights on 18 pressure levels from 1000 to 0.4 hPa at NDACC stations, which include DDU. Rephrased line **185** for better clarity into "provides direct extraction of the temperature product above DDU"

https://www-air.larc.nasa.gov/pub/NDACC/PUBLIC/meta/ncep/ncep_2022.pdf

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?

NCEP provides an error estimation of the temperatures at DDU of 2.5 K from the surface to 50 hPa, then 3.5 K at 30 hPa and 10 hPa and more than 5 K above 5 hPa. The errors on the geopotential heights provided at DDU range between 16 m near the surface to 150 m at 5hPa.

ERA5 uncertainties are discussed at:
https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation#ERA5:datadocumentation-Accuracyanduncertainty but no specific uncertainty field is provided for in the ERA5 product.

For Antarctica, ERA5 validation in the Southern Antarctic suggests a warm bias of 0.14 °C, with a significant improvement compared to the ERA5 predecessor, ERA-Interim, in particular at high altitudes (Tetzner, D.; Thomas, E.; Allen, C. A Validation of ERA5 Reanalysis Data in the Southern Antarctic Peninsula—Ellsworth Land Region, and Its Implications for Ice Core Studies. *Geosciences* 2019, *9*, 289, doi:10.3390/geosciences9070289).

Our study does not aim at validation nor assessing the quality of the reanalysis. Moreover, since it is very complicated to discuss these biases and how they evolve temporally/seasonally, and in particular over the location of DDU, we choose not to discuss the reanalysis uncertainties in the new version of the manuscript. More information on the intercomparison between NCEP and ERA5 is heavily listed in comments 74 and 79 hereafter.

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.
This paragraph has been reworked, we hope the reviewer will find it clearer. "dynamic" here was a poor wording and was removed. The idea is that the time averaging does not split the measurement sessions blindly, but takes the atmospheric situations sounded into account:

The purpose of the pre-processing step is to avoid blending in different atmospheric scenes when choosing the integration time windows. Averaging 30 min of measurements which contain 15 min of measurements of a cloud and 15 min of clear air will result in a somewhat smoothed signature that does not clearly point to any physical observational reality. To avoid this effect, we first perform a first-order inversion assuming no particulate extinction on every 15 minutes file. The resulting approximated scattering ratio profiles are then flagged to state on aerosol / cloud presence. Doing so, if for instance a measurement session captures an atmospheric change like the appearance / disappearance of a cloud, we are able to cut the session at the time of this change to avoid blending in different situations. This is what we call a homogeneous atmospheric scene: a time window during which the stratospheric column sounded by the lidar is globally stable signal-wise.

After this step, the raw 3 minutes files are integrated following the results of the pre-processing algorithm and proper signal inversion is performed. Once the R_T and δ_{aer} computed, a peak detection algorithm is applied, identifying the peaks in the signal. Peaks identification is a typical feature in signal processing with many available implementations, the challenge being to separate outliers using input parameters not specific to our instrument. Peaks detected in the R_T and δ_{aer} profiles correspond to scattering layers. Once the layers have been identified, the relevant parameters are computed for each layer. As mentioned in the manuscript, there are three sets of such parameters, which correspond to the three classifications considered in this study. Therefore, the result of this detection algorithm consists of three sets of layers and their optical parameters, corresponding to the three classifications B05, P11 and P18. The altitude and [R_T, δ_{aer}] of the layers are the same in each of these three sets, but their interpretation changes according the scheme used.

This is performed on every single measurement, the only purpose of the pre-processing phase is to select the integration time windows that split measurement sessions. No individual lidar profile is discarded in the process. Of course, measurements that were flagged as clear sky situations in the pre processing will most likely not result in any scattering layer identification as the optical parameters will not overcome the background thresholds. However, since the pre-processing phase is based on an approximated inversion, we consider it safer not to exclude them from the layer detection procedure.

Above statements are now included as edition in lines **188-194**: "Before the inversion procedure, lidar data are first pre-processed to adequately set the integration window of the individual 3 minutes raw files ensuring homogeneity of the lidar scene, either being dominated by clear-air or aerosol/cloud presence. To do so, a preliminary inversion assuming no particulate extinction is performed on fifteen minutes blocks to derive a non-corrected for extinction (nce) scattering ratio. A clear sky or Aerosol/cloud presence tag is applied to this preliminary product. Then, the 3-minutes raw files are summed accordingly. This pre-processing avoids the spatio-temporal smoothing

necessarily induced by blending in clear-air before or after any cloud detection and leads to a better Signal to Noise Ratio (SNR)."

50) Line 144: "assuming no particular extinction" – do you mean "assuming no particulate extinction"?
Yes, mistake corrected. Edited "no particulate extinction" line **190**

51) Lines 155-169: As mentioned above, I suggest you move these lidar parameter definitions up to Section 2.1
This point was addressed in comment 33).

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.
The point is actually close to the goal of the paper, and the introduction has been reworked to make it clearer ; besides, the sentence pointed out by the reviewer is indeed misleading. We acknowledge that the location of DDU and the associated dataset may not be be representative of the whole antarctic continent and therefore not suited to validate the full PSC spectrum covered by the classification schemes, especially regarding the larger ENH and ICE particles. In turn, we find it meaningfull to consider different schemes built on different optical properties to enrich the analysis our data. The paper aims at providing the community with a PSC distribution above the station. The consistency of this distribution may only be assessed from different schemes, and this work remains to our knowledge rarely made, the most recent one being Achtert et al. (2014) and applied to the Arctic.

The sentence pointed out by the reviewer was modified to present the purpose of this section more clearly, edited lines **228-230**: "The purpose of this section is notably to provide the community with a PSC distribution which consistency has been checked using different observational parameters and thresholds. Applying different classification schemes emphasizes the variability and in turn, consistency, of our distribution."

53) Lines 176-182: Much of this was already discussed in the Introduction.
It is indeed repetitive, and was replaced by the following sentence, lines **230-232**: "As mentioned in introduction, the classifications B05 and P11 are considered here following the conclusion of Achtert et al (2014), to which we had P18, the update of P11 published in 2018."

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)?
It is a mistake, we are indeed talking about 2 orders of magnitude. Edited "two orders of magnitude" line **276**

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.
Following suggestions of all reviewers, a comparison of CALIOP data in the DDU vicinity and DDU lidar was added to the study. Therefore, we extracted the $R_{T,ICE}$ threshold values used by

CALIOP around DDU and decided to use them in our analysis fueled by the P18 classification. The manuscript was edited accordingly, lines **281-283**:

"We choose to use the $R_{T,ICE}$ values provided by CALIOP in the PSC Mask v2 product. For each DDU lidar measurement, the $R_{T,ICE}$ value was taken from the closest CALIOP profile in both time and space, within the 100 km radius area centered on the station."

56) Line 235: Suggest rewording this to read: "The wave ice category defined in P11 and P18 was ignored in this study ..."
Edited as suggested: "The wave ice category defined in P11 and P18" line **284**

57) Line 236: verb tense doesn't match- "... published classifications ... features" should be "...published classifications ... feature"
Typo corrected line **287**, "feature".

58) Line 242: Suggest changing "disequilibrated" to "non-equilibrium"
Corrected, "non-equilibrium" line **292**

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.
We thank the reviewer for noting this mistake. Reading Tesche et al. (2021), we actually missed that the Antarctic comparison is only based on 2012 and 2015 while the Arctic one is based on all winters from 2006 to 2018. To address this important mistake, we extracted the CALIOP measurements around DDU from 2007 to 2020 and produced a new figure and edited the analysis accordingly.

To properly account for this new content, a section 2.2 was added lines **140-147** to introduce CALIOP. A section 3.2, lines **202-223**, was also added to present the PSC detection method of CALIOP compared to the one we use. A paragraph was added in the beginning of section 4.1., lines **296-309**, to detail the statistics of CALIOP overpasses above DDU. The discussion of figure 3 was adapted adequately.

In the reviewed PSC distribution, we count the PSC layers for each type. Introducing a comparison with CALIOP measurements leads us to adapt our counting approach. Since CALIOP sorts each vertical bin separately (with a coherence criterium), it takes the geometrical thickness of PSC into account. This is relevant and we modified our method accordingly. Therefore, we now take into account each vertical bin of the identified PSC layer as explained in section 3.2. This explains the change in the distribution in figure 3 as compared to the reviewed version.
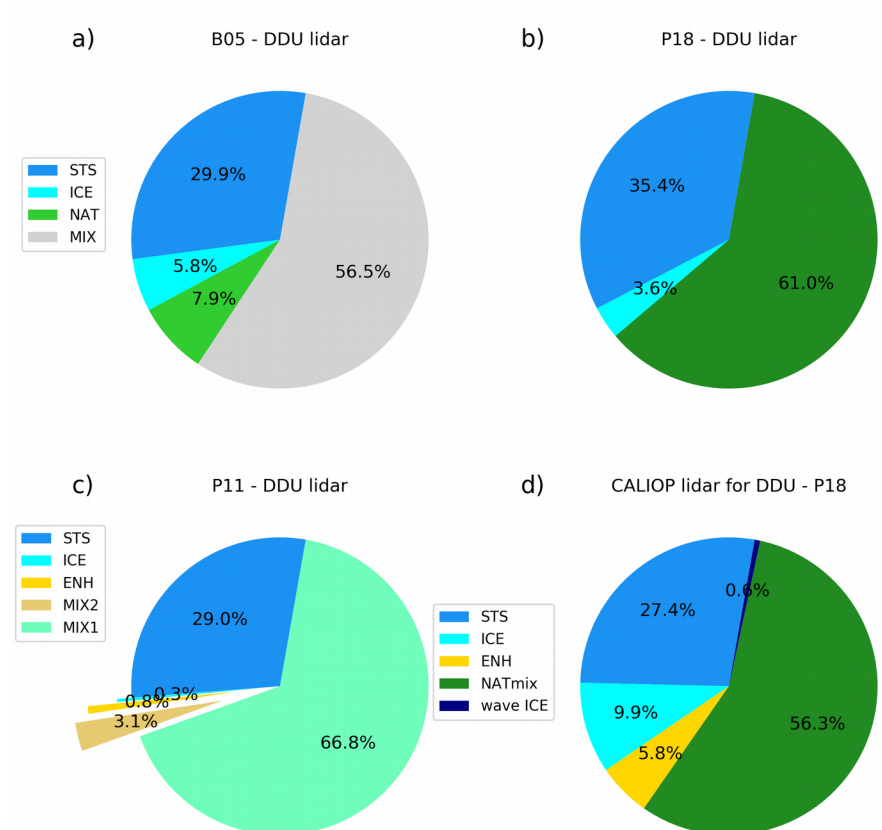
Here is the new version of Figure 3:

Figure 3: PSC types distribution observed at DDU for the three considered classifications: B05 (a), P18 (b), P11 (c) and observed by CALIOP extracted around DDU using P18 scheme (d).

Edited lines **198**: "... is attributed to each vertical bin of the layer."

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.
We indeed extracted CALIOP measurements at DDU as explained before. Doing so, we could identify that ICE PSC have only been detected by CALIOP at DDU on 19 different days between 2007 and 2020. This sample seems too low to build statistics. We can however say that DDU lidar was rarely operated on these days (only 5 out of these 19 days), suggesting a possible important tropospheric cover or bad weather condition hindering operations. We now mention this in the manuscript.

Edited lines **335-338**: "Between 2007 and 2020, CALIOP detected ICE PSC above DDU on 19 different days, out of which 4 correspond to DDU measurements, suggesting a possible important tropospheric cover or bad weather condition hindering operations. However, we do not consider this small sample robust enough to support the analysis."

61) Lines 280-283: These two sentences are not clear- not sure what you're trying to say.
Our point here considers the depolarization threshold used in B05 (10%) along with the few NAT clouds identified at DDU using this threshold: we consider the latter number to be too conservative given the expected abundances of NAT clouds. As compared to other classifications which include "NAT mixtures" (MIX1 and MIX2 in P11 and NATmix in P18), the definition of NAT may be more restrictive in B05. This is what was implied by "pure NAT clouds": we can consider that the NAT

clouds identified with B05 are PSCs entirely composed of – or at least highly dominated by - NAT particles, and not mixed with significant amounts of STS droplets.
This was edited to make it clearer, lines **360-363**:
"Given the 10% depolarization threshold and the relatively low amount of NAT clouds identified by B05, we consider that B05 classifies as NAT the PSC that are only composed of, or highly dominated by NAT particles. Whereas P11 and P18 separate NAT mixtures into MIX1, MIX2 and NATmix which may include a significant share of STS droplets."

62) Lines 285-286: Why would optical properties be dependent on latitude?
Sentence actually needs to be rephrased ; we do not imply some direct correlation between latitude and optical properties, but rather that some range of optical properties need thermodynamical conditions rarely met at lower antarctic latitudes : ICE PSC persistence is related to temperature remaining below T_{ICE}, and out point is simply to state that optically equilibrated persistent ICE PSC layers are rarer above DDU than above stations at higher latitudes.

Edited in the manuscript to be clearer, lines **365-367**:
"The relative low number of ICE events we report relates to the fact ICE PSC fields above DDU are more unstable that those remaining deep inside the vortex. The optical properties of the ICE clouds observed at DDU are thus expected to be closer ...."

63) Line 290: "barely never" – that phrase makes no sense. Do you mean "barely ever"? Probably would be better to just say "rarely detected."
Corrected as suggested "rarely" line **370**

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?
Figure 5 illustrates several points that sound consistent to us. First, it shows the altitude distribution of each PSC type. It offers another point of view of the different distributions produced by the three classifications shown in Figure 3.

Considering this comment and the following concerning PSC formation temperature thresholds, we indeed made the approximation of early winter conditions. The idea was not to use MLS measurements at DDU, just like the $R_{T,ICE}$ threshold in P18. Following the reviewer's comments we now do calculate T_{NAT}, T_{STS} and T_{ICE} using the closest MLS H_2O (and HNO_3 when necessary) measurement available. To include this change in figure 5, we needed to change its structure : the figure now presents the temperature relative to the relevant threshold (T_{NAT}, T_{STS} or T_{ICE}) depending on the type considered. For each PSC measurement, we calculated the threshold temperature corresponding to the daily abundance of H_2O and HNO_3. Still, this approach doesn't account for the kinetics of PSC formation, and embeds the relatively large uncertainties of MLS measurements at PSC altitudes. Thus, the temperature shifts observed mainly for the STS class are completely acceptable to us, especially considering STS formation temperature is apart from the NAT and ICE one, as it is not associated to a discrete physical phase transition but to a continuous chemical composition change within the droplet.

The analysis of Figure 5 has been adapted to the new layout. The new version of Figure 5 is shown below:
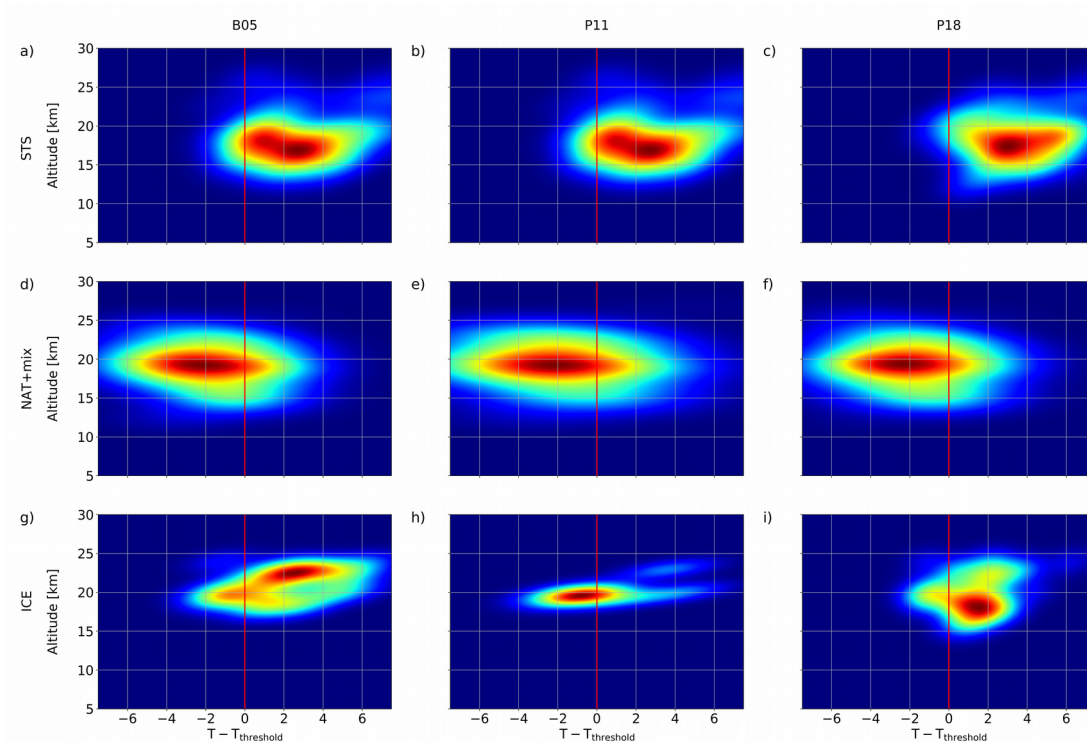

Figure 5: Distribution of PSC detection at DDU from 2007 to 2020 as a function of altitude and temperature relative to the relevant threshold, T_{STS}, T_{NAT} or T_{ICE}. Lines correspond to STS, NAT + mixtures and ICE from top to bottom. Columns correspond to the classification schemes B05, P11 and P18 from left to right.

Lines **380-381**: "For each PSC measurement, MLS H_2O and HNO_3 concentrations are used to compute the relevant threshold temperatures."

Lines **382-383**: Figure 5 presents the distribution of PSC measurements as a function of altitude and temperature relative to the relevant type threshold."

Lines **384-389**: "... Temperature is not a variable in our PSC detection method, yet we note that most NAT+mix measurements are below the T_{NAT} threshold within expected uncertainties related to MLS and ERA5 data. Considering STS, it appears that temperatures are above the threshold for all schemes. This is partly expected as T_{STS} is not associated to a discrete physical phase transition but to a continuous chemical composition change within the droplet. Finally Fig. 5 highlights again that the major difference among classification schemes concerns the ICE category as the three patterns of Fig. 5g, h and i show very different shapes."

Lines **390-391**: "It appears however that P11 classifies few PSC as ICE but those are in the adequate temperature range while B05 and P18 sort most of the ICE PSC above T_{ICE}."

Lines **398-399**: "Also, the time of lidar PSC measurement does not exactly match ERA5 data and MLS H_2O and HNO_3 measurements which necessarily generates uncertainty."

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?

This example illustrates some features of PSC fields and bring some fine small-scale considerations as compared to the global point of view of classification schemes.

For example, it highlights the stratified layering mentioned earlier in the paper and in the review. It also highlights the variability as compared to the fixed thresholds of PSC classifications.

This section is not included to be representative of 14 years of measurements, it is here to illustrate several points mentioned earlier in the article, and also aims at showing that groundbased lidars have opportunities to access to the small scale structure and dynamics of PSC fields.

A statistical analysis could be made on the interannual variability, which we did by providing the annual variability in the trend analysis. A interannual variability of the PSC type distribution would use smaller samples not robust enough to produce reliable statistics. Moreover, the low amount of CALIOP overpasses above DDU would further harm the statistics. The current version of the paper provides statistics on CALIOP PSC observation above DDU, highlighting the relatively low PSC occurences at DDU as compared to higher latitude stations.

Even if we don't consider yearly PSC type distribution, the interannual variability of PSC occurences is displayed in Figure 9 as red triangles marks. Besides, the new version of Figure 9 now includes the number of CALIOP PSC days at DDU, merely above 10 per year.

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.

The mistake in figure numbering was corrected. Edited "Figure A1" line **406**.

We found interesting to mention the agreement with Reprobus for this observation. However, since we did not discuss further this model, the figure was only included in annex. To us, there is no sufficient analysis supporting putting this plot in the core of the paper, especially considering the limitation of PSC modelling in REPROBUS. It is purely presented as mesoscale context.

67) Line 324: Again think you mean Figure A1

Again corrected, line **410** "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."

The statement was indeed too strong. It was edited accordingly, lines **411-412**: "The model produces PSC presence at the 435 K and 475 K levels, and no PSC at the 550 K level above DDU, which is 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."

Rephrased in a safer manner as suggested, line **419**: "has not been accounted for".

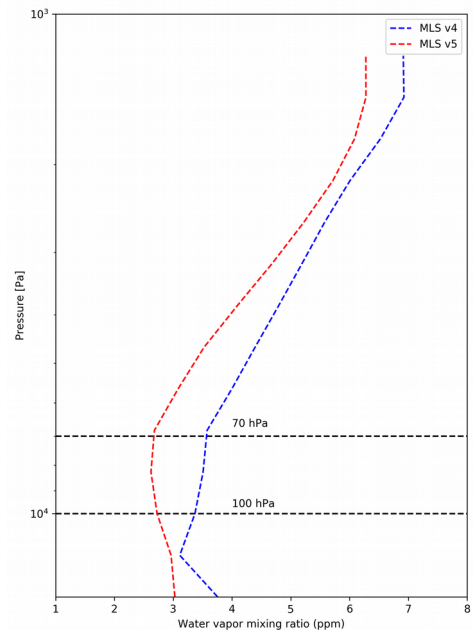70) Line 334: RT and RICE should be written as R_T and R_{ICE}

The indices have been corrected, line **420-421**.

71) Figure 5 discussion: Are the T_{ICE} thresholds based on 5 ppmv H_2O? This is not likely representative of 28 August when the stratosphere may be severely dehydrated. How would this change your interpretation?
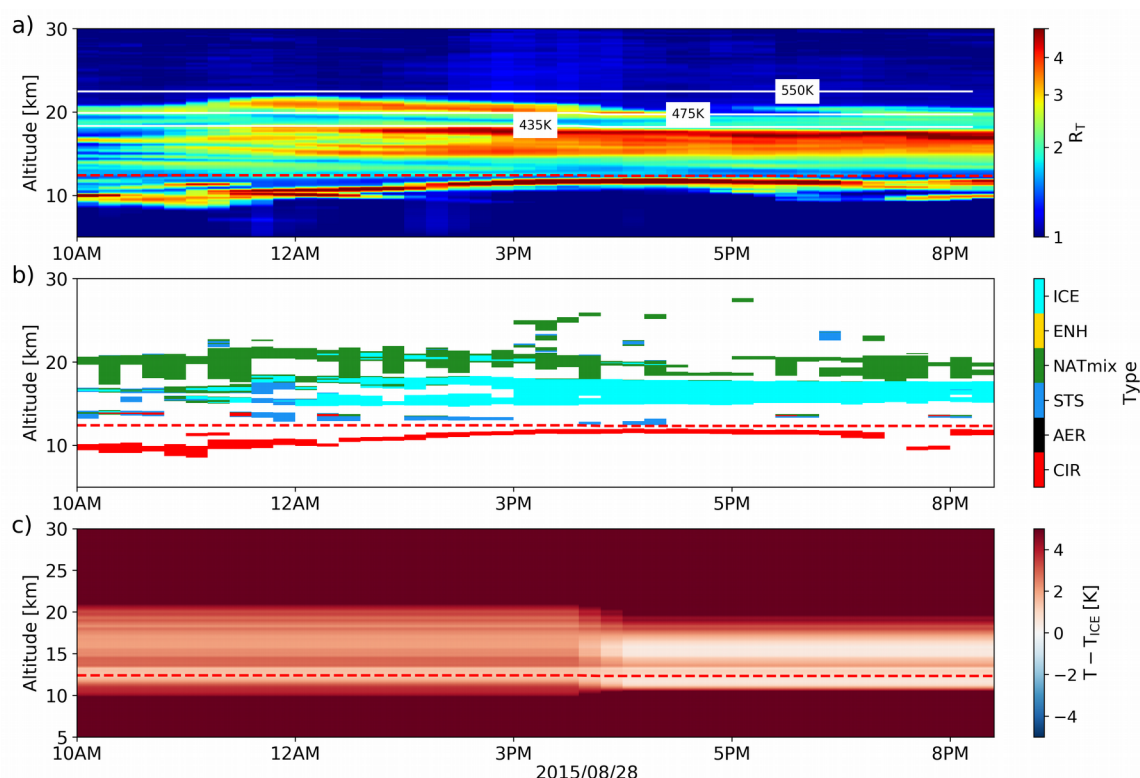
Figure 6: 532 nm backscatter ratio of lidar measurements obtained at DDU on the 2015/08/28 (a). The corresponding PSC types according to P18 classification scheme (b) and ERA5 temperatures at DDU as compared to the ICE formation threshold T_{ICE}, calculated from the MLS H_2O of the day (c). The red dashed line indicates the dynamic tropopause computed from ERA5 data.

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?

They are indeed based on the difference between NCEP, ERA5 and IASI at the DDU location and the radiosonde at its true location. In other words, the radiosonde drift has not been taken in account here. We think the comparison is still relevant and we extensively detailed this choice in our response to the comment 79.

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?

The issue on the temperature bias caused by the radiosonde drift is extensively discussed later on in comment 79.

As for the model intercomparison, we do not claim the ERA5 products are better than the latest NCEP reanalyses. Historically, ECMWF stratospheric meteorological products tended to be a bit more accurate than NCEP products (see references below), notably for products relevant to our study (i.e. stratospheric Antarctic temperatures).

However, despite the multiple publications from the SPARC Reanalysis Intercomparison Project (S-RIP) (such as the special issue on this topic in Earth System Science data: https://essd.copernicus.org/articles/special_issue10_829.html), it is not very obvious that the latest ECMWF reanalyses, ERA5, still have an edge on the NCEP products.

When compared to the high vertical resolution Global Navigation Satellite System radio occultation (GNSS RO) data, ERA5 shows obvious improvements in temperature data compared to ERA-I and also a slightly better agreement with GNSS RO measurements than MERRA2. For the 2007–2017 period, ERA5 shows the best agreement with observations, while the other two reanalyses (ERA-I, MERRA2) slightly warm biased. Clearly, ERA5 reanalyses represent a substantial improvement on ERA-I reanalyses which were already widely used (Hoffman et al., 2019).

We have not found a single publication demonstrating the better quality of the NCEP products compared to ERA5.

https://rmets-onlinelibrary-wiley-com.insu.bib.cnrs.fr/doi/abs/10.1256/qj.03.76
Nevertheless, it is found that NCEP/NCAR reanalyses tend to be slightly warmer (0.8 K) than the observations, while the converse is true for ECMWF analyses (−0.3 K). Finally, trajectory comparisons are performed. It is found that trajectories built with ECMWF winds are more accurate than those built with NCEP/NCAR winds.

https://agupubs-onlinelibrary-wiley-com.insu.bib.cnrs.fr/doi/full/10.1029/2008JD010116
This article compares the temperature, zonal, and meridional velocities issued by the 50-year National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP-NCAR) Reanalysis (NN50) and European Centre for Medium-range Weather Forecasts (ECMWF) operational analyses with independent observations collected during the Vorcore superpressure balloon campaign. The ECMWF analyses are found to be more accurate than the NN50 reanalyses. In particular, an overall warm bias in the polar Southern Hemisphere lower stratosphere is found in NN50 (+1.51 K), while a cold bias is found using ECMWF analyses (−0.42 K).

The results of these comparisons indicate that NN50 tends to be a few degrees colder than the observations in the SH subpolar latitudes, while ERA-40 is less hit by this cold-pole issue

In 2000 the standard deviations of ECMWF, MO, and DAO with respect to the measured temperatures range from 1.0 to 1.3 K, whereas NCEP and REA have substantially larger errors. In 1999 the flights took place during a major warming, and all operational models had large standard deviations and substantial biases. Preoperational versions of the new ECMWF model with increased stratospheric resolution and assimilation of the advanced microwave sounding unit, which none of the other models assimilated, show small biases and standard deviations.

75) Lines 388: This statement should appear in the Introduction and help define the focus of the paper!
This statement was moved in introduction as suggested and rephrased accordingly in this section in order to avoid repetition.

Moved to introduction, lines **82-84**: "Two of the major roles of a ground station are to perform process studies and establish decadal trends. Such trends are highly valuable because they reflect the evolution of the stratosphere, in terms of temperature and chemical compositions. "

Edited line **475**: "As mentioned in introduction, ground stations are key to the establishment of decadal trends."

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?
As discussed in a previous comment (74), NCEP is considered less accurate at DDU than ERA5. It still does not make the comparison proposed in Figure 9 useless. We are not using NCEP instead of ERA5, we are using both. And we modified Figure 9 to also include IASI as this was suggested by another reviewer. The inclusion of NCEP here adds valuable information in that, despite the discrepancies between datasets highlighted by figure 8, the two model and the spaceborne observationally derived trends are globally consistent.

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?
The original version of the figure indeed included a constant T_{NAT} value calculated from a 10 ppb value for HNO_3 and 5 ppm vor H_2O. Daily MLS measurements were extracted around DDU and the T_{NAT} values are now calculated based on these time dependent values. The trend and the corresponding figure have therefore been changed. The trend shifts to -5.7 to -4.6 PSC days per year per decade without significant change on the analysis.
It is important to note that taking the stratospheric denitrification into account tends to decrease the computed T_{NAT} values. As an expected result, the ΔT criteria adujsted to our lidar measurements is now -1 K and not -2 K as in the initial version of the manuscript. Figure 9 and the associated discussion have been edited on lines **490** and **492** and in the caption of Figure 9.

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.
Now that we have extracted CALIOP PSC Mask v2 product around DDU, we can include such a comparison. As it was highlighted in the results of the comparison of CALIOP and DDU lidar

measurements, there are few CALIOP profiles featuring PSC detection available at DDU, so that computing a trend per year is not reliable. The number of days with CALIOP PSC detection at DDU, per year, is now featured in figure 9 as black triangles. We find on average 6 PSC days per year detected by CALIOP around DDU, with a minimum of 1 (2012, 2014) and a maximum of 11 days (2009). We note that this would produce a slightly negative trend of -1.1 PSC days per decade, but we consider that building a trend on so few PSC detections is not significative as discussed in a previous comment.

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?
First, let us state on the radiosonde drift: radiosonde temperature profiles are not used in the lidar inversion because of the low burst altitude in winter. We mostly use them to further investigate on borderline situations such as discussing the type of a PSC detected or the temperature related to an aerosol layer.

When doing so, it is important to keep in mind that the radiosondes can experience significant drift, especially in winter, and therefore not reflect the temperature of the atmospheric column sounded by the lidar. However, this does not mean that these radiosondes are not globally representative of the climate aroud DDU and that they cannot be used to evaluate the relevance of temperature datasets at DDU, the drift remaining below the 1° pixel resolution in model grids. The plot below shows the mean distance of the radiosondes to DDU for each 1 km interval, for the whole year, summer and winter.

There are two main reasons why we consider that the comparison of the temperature datasets with the radiosondes is relevant despite the point we made about their drift.

The location of DDU is 66.6°S – 140°E. IASI extraction area around DDU is set to 66.8°S – 66.5°S and 139.7°E – 140.3°E, which correspond to approximately a rectangle of 26,5 x 33,4 km.

The mesh of the ERA5 product used in our study is 0.25° x 0.25°, which is an area of approximately 27.8 x 11.1 km at DDU.

NCEP provides temperature and altitude profiles for DDU station, at the location 67°S – 140°E. The real location of the station being 66,66°S – 140°E, NCEP interpolation is approximately 37,8 km from the true lidar location.

Therefore, considering that up to an altitude of 15 km the radiosondes are on average closer than 50 km of DDU (see figure attached), and considering the distances at stakes for the extraction of IASI measurements or the sizes of NCEP and ERA5 meshes, it is reasonable to say that these datasets are representative of the same area. When comparing NCEP, ERA5 and IASI to these radiosondes as shown on figure 7, even when only focusing on altitudes below 15 km, it appears than NCEP is less accurate than ERA5 and IASI at DDU.
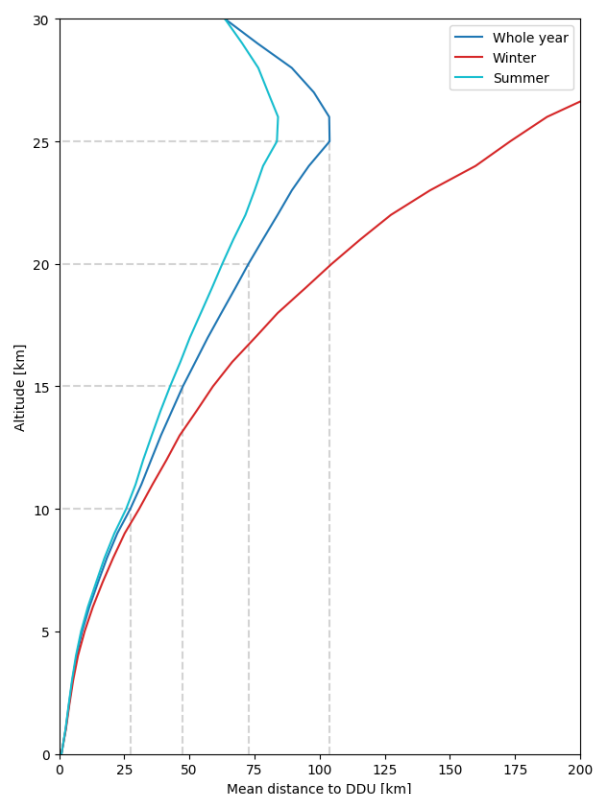
The other reason is more pratical. In order to account for the drift of the radiosondes in figure 8, we would need to adjust NCEP, ERA5 and IASI at the true location of the radiosonde every day. While it is realistic with the 4xdaily ERA5 and its 0.25°x0.25° grid, it would raise issues for IASI and NCEP. First, the NCEP product used in this study is provided already interpolated at DDU location, so we would need to use another product to access to temperatures around DDU.

Depending on the location of DDU in the field of view of IASI, taking the drift of the radiosonde in account would sometimes imply to produce a temperature profile composed of different orbits and therefore different times of measurement, and then compare it the radiosonde. This would make no sense and would in many cases be worse than ignoring the drift of the radiosonde in the comparisons.

Finally, NCEP interpolation at NDACC station is provided daily, while the ERA5 product we interpolate at DDU is available 4xdaily, which is of course an advantage to be as close as possible to the lidar measurements.

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?
The discussion of this figure may not be clear enough and related to the way the figure is built. For each PSC detection at DDU, we computed the total stratospheric domain occupied by PSC layers, i.e. total PSC thickness and we also computed the height of the stratospheric domain satisfying $T < T_{NAT}$ on this very day. Then, the distribution of both values computed for the whole dataset were plotted. So the highlight of the figure is that, while PSC thickness is approximatively equally distributed between 1 and 7 km, the stratospheric domain under T_{NAT} when PSC are detected is significantly higher, often larger than 10 km. The point is to show that PSC layers do not fill the whole stratospheric domain satisfying $T < T_{NAT}$, because PSC volume were sometimes evaluated this way.

We consider that this figure provides interesting information on the thickness of PSC layers as compared to their potential "domain of existence".

Since this figure was not clearly understood by two reviewers, we adapted it to be more straight forward. For each day with a PSC detection, we calculated the stratospheric range satisfying the condition $T < T_{NAT}$. T_{NAT} was calculated with the daily MLS H_2O and HNO_3 measurements. Let us call this range H_{NAT}, expressed in km. We also calculated the geometrical thickness of the PSC detected, called H_{PSC}. Then, we computed the difference $H_{NAT} - H_{PSC}$: it represents the stratospheric range satisfying $T < T_{NAT}$ unoccupied by PSC layers. The distribution of this difference, for all PSC detections at DDU from 2007 to 2020 is plotted in Figure 10. It shows that PSC do not occupy all the stratospheric volume satisfying $T < T_{NAT}$ so that estimating the first quantity by computing the second is not a satisfying assumption. On this figure are only represented the PSC measurements for which ERA5 temperatures dropped below T_{NAT} otherwise it would have featured negative values. Such cases represent 29% of all PSC detection, they can be caused by the approximation in the calculation of T_{NAT}, the uncertainty of ERA5 or an outlier in our PSC detection method.

Here is the new version of Figure 10:

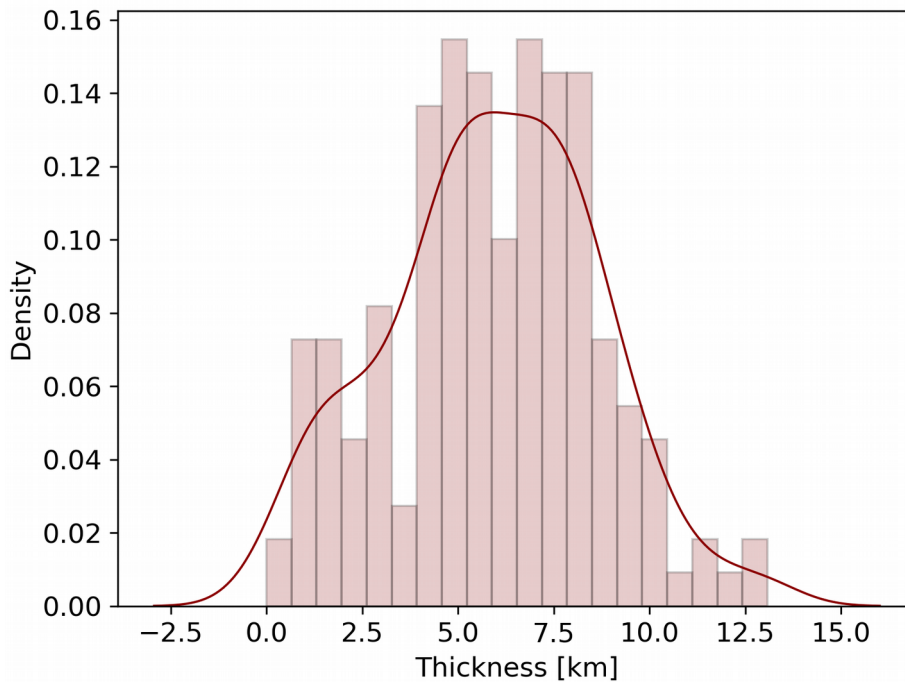Figure 10: Distribution of the thickness of the stratospheric domain satisfying T<T$_{NAT}$ unoccupied by PSC layers, in km, on days when a PSC is detected at DDU.

81) Section 5 Conclusions: No specific issues at this point. Will reserve comment for the anticipated revised version.

References:

Ansmann, A., Ohneiser, K., Chudnovsky, A., Knopf, D. A., Eloranta, E. W., Villanueva, D., Seifert, P., Radenz, M., Barja, B., Zamorano, F., Jimenez, C., Engelmann, R., Baars, H., Griesche, H., Hofer, J., Althausen, D., and Wandinger, U.: Ozone depletion in the Arctic and Antarctic stratosphere induced by wildfire smoke, Atmospheric Chemistry and Physics, 22, 11 701–11 726, https://doi.org/10.5194/acp-22-11701-2022, 2022.

Rieger, L. A., Randel, W. J., Bourassa, A. E., and Solomon, S.: Stratospheric Temperature and Ozone Anomalies Associated With the 2020 Australian New Year Fires, Geophysical Research Letters, 48, e2021GL095 898, https://doi.org/https://doi.org/10.1029/2021GL095898, e2021GL095898 2021GL095898, 2021.

Snels, M., Scoccione, A., Di Liberto, L., Colao, F., Pitts, M., Poole, L., Deshler, T., Cairo, F., Cagnazzo, C., and Fierli, F.: 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, https://doi.org/10.5194/acp-19-955-2019, 2019.

Snels, M., Colao, F., Cairo, F., Shuli, I., Scoccione, A., De Muro, M., Pitts, M., Poole, L., and Di Liberto, L.: 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, https://doi.org/10.5194/acp-21-2165-2021, 2021.

Stone, K. A., Solomon, S., Kinnison, D. E., and Mills, M. J.: On Recent Large Antarctic Ozone Holes and Ozone Recovery Metrics, Geophysical Research Letters, 48, e2021GL095 232,

https://doi.org/https://doi.org/10.1029/2021GL095232, e2021GL095232 2021GL095232, 2021.