We would like to thank the Referee for the valuable input. Please find out point-by-point reply below. Referee comments are given in black, our answers are given in blue.

Anonymous Referee #1

The paper addresses the representativeness of ground-based lidar measurements in the Polar regions with respect to CALIOP (and MIPAS) observations of polar stratospheric clouds. The main conclusion of the paper is the identification of the best sites for PSC observation. To my opinion, the title is not adequately describing the main goal of this work. I would suggest something like "How to find the best locations for ground-based PSC observations", which better expresses the conclusions and recommendations of the authors.

We thank the Referee for this suggestion. It is indeed an outcome of the paper to define the best locations for ground-based PSC observations. Nevertheless, we think that the current title is appropriate as we show results for the entire Arctic and Antarctic, respectively, in Figures 2, 3, 5, and 6. These maps provide information that goes further than just the locations of existing research stations.

The comparison of the two CALIOP datasets (troposphere and PSC v2) and the ground-based lidar observations might produce many interesting results. The paper does not fully explore the potential of this method and also is not considering possible biases due to the different measurement protocols of CALIOP and ground-based lidars. It would be useful to specify the different categories of ground-based lidars; those measuring in a continuous mode, others "randomly" and still others in a CALIOP-synchronous mode". The authors should also explain that CALIOP is NOT a continuous mode lidar at a certain location, but has overpass frequencies in the order of days at specific local times. This might cause a bias in the statistics.

We are sorry that our description was not as clear as intended. The scope of this study is to explore the effect of tropospheric cloudiness on what would be observed from ground. The basic assumption is that the CALIPSO lidar can observe all PSCs along its laser beam while tropospheric clouds might attenuate the laser beam before it can reach PSC altitudes. In other words, we look at the same observation from two directions, i.e. from ground and space. In terms of the measurement protocol of a ground station, this implies CALIPSO-synchronous mode. We then separate between scenarios in which measurement at ground are performed (i) during each CALIPSO overpass and (ii) only one third of all CALIPSO overpasses. The first scenario can be realised both with a continuously operating lidar and with a system that is operating during each CALIPSO overpass with downtime in between but only if there is no interference by tropospheric clouds or measurement-inhibiting factors such as maintenance, downtime, or operator availability. The second scenario also refers to CALIPSO-synchronous measurements with the caveat that interfering factors reduce the number of observations to one third. This latter scenario is much more realistic for most polar lidar stations.

We agree that the labelling of a continuously operating lidar and a manually operated lidar was misleading. Accounting for this and the fact that we really only consider the CALIPSO-synchronous scenario, we have dropped the reference to continuously operating and manually operated instruments. We have also revised the text in Section 2.4 to:

"*The matched observations of tropospheric and stratospheric clouds allow for a direct comparison of individual PSC profiles as well as long-term PSC statistics as seen from ground and space independent of the considered instruments. Specifically, the same profile can be evaluated from two perspectives, i.e. from space as well as from the point of view of a ground-based instrument. In that context, the latter perspective translates to a CALIPSO-synchronous measurement protocol at a ground station. True PSC statistics unaffected by tropospheric cloudiness, i.e. during all-sky conditions, at a certain*

location can only be obtained with a spaceborne lidar. In contrast, filtering with respect to tropospheric cloudiness is applied to emulate the likely conditions for meaningful ground-based PSC measurements in the CALIPSO data set. Specifically, we assume that a ground-based lidar would only provide meaningful results during conditions with no clouds or only transparent clouds that would not already attenuate the laser beam before it can reach PSC altitudes. This is referred to as the ground-based view of the CALIPSO data set. It provides sampling that is dependent on the CALIPSO return rate and must not be confused with actual ground-based measurements that can provide localised PSC observations in the time range from hours to weeks.

We subsequently separate the ground-based view of the CALIPSO data set into two scenarios for which (i) all cases of the ground-based view are considered and (ii) one third of the profiles of the ground-based view was randomly selected. The first scenario corresponds either to a continuously operating lidar or a manually operated system that is active during every single CALIPSO overpass with possible downtime in between without any interference by tropospheric clouds or measurement-inhibiting factors. The second scenario also refers to CALIPSO-synchronous measurements with the caveat that interfering factors reduce the number of measured lidar profiles to one third of what would ideally be possible. This latter scenario is much more realistic as (i) most ground-based lidar instruments are operated manually and on campaign basis, (ii) the decision to start a measurement, i.e. the assessment of tropospheric cloudiness, is made subjectively by the operator, and (iii) infrastructural challenges (e.g. system downtime, logistical problems, and lack of personnel) affect the operation of a ground-based lidar at a remote location and under harsh conditions.

To assess the representativeness of ground-based PSC measurements, PSC statistics are obtained for boxes of 2° latitude by 2° longitude around the sites in Figure 1 and Table 1."

Having at disposition both data sets the authors might also explore the possible correlation between tropospheric cloudiness and PSC occurrence (as they mention in lines 240-247).

The investigation of the connection between tropospheric and stratospheric cloudiness has actually been our motivation from the outset. The present study turned out to be a by-product of this work and we decided to publish it first as makes for a nice stand-alone publication.

They also might quantify the bias introduced by prohibitive meteorological conditions, such as cloud cover in the ground-based dataset, by comparing the PSC occurrence, as observed by CALIOP, with and without cloud cover. I suppose that this could be easily done.

This is actually the scope of the manuscript. We use the matched CALIPSO observations of tropospheric and stratospheric cloudiness to show what PSC statistics look like during (i) all-sky conditions (spaceborne view, not possible with ground-based instruments), (ii) situations with tropospheric cloudiness that would still enable PSC observations from ground (view of a ground-based lidar with CALIPSO-synchronous measurement protocol operated during every single CALIPSO overpass), and (iii) situations in which CALIPSO-synchronous operation of a ground-based instrument is affected by cloudiness and other measurement-inhibiting factors. We have revised the description of the data analysis in Section 2.4, the caption of Figure 4, and the discussion of Figures 4 for clarity. We have also made revisions throughout the text to clearly state that the purpose of this work is exactly to quantify the bias introduced by prohibitive meteorological conditions, though we refer to them simply as tropospheric cloudiness.

An important flaw of the paper is that they apparently are not aware of the fact that a lidar observatory is active at Concordia station since 2014 (see e.g Snels, ACPD 2020 and

https://tmf.jpl.nasa.gov/testLidar/NDACC_LWG/sites/dome_c.html). This is particularly relevant, since the authors recommend Concordia as one of the best sites to perform PSC observations.

We thank the Referee for making us aware of this publication and the measurements at Concordia station. The paper has been added and the Figures and discussion have been revised to account for the existence of PSC measurements at Concordia (see also replies below).

The authors consider the CALIOP observations as a reference system for the ground-based lidar. When they speak about representativeness they refer to the agreement of the statistics of the ground-based lidar measurements with respect to the CALIOP observations. This is generally speaking an acceptable concept, but there are some caveats. CALIPSO is performing 14-15 orbits per day, which means that the orbits have a separation in longitude of about 180/15 = 12 degrees (we have ascending and descending overpasses). At a latitude of 70(80) degrees. 12 degrees of longitude means 450 (225) km of distance between successive overpasses. The authors use boxes of 2 x 2 degrees lat-lon boxes to do their statistics, this means that several days are needed to "fill the boxes". Experience shows that tropospheric clouds and PSCs are not constant over days, often they change during the day. The CALIOP overpasses in a box occur at fixed local times and thus are biased wrt to the random ground-based observations. Synchronized ground-based observations eliminate this bias. If one considers only average statistics, one should take into account the biases present in the comparison of ground-based lidar observations wrt to CALIOP, due to the different measurement times. Some stations (McMurdo in the past, Concordia in the present, maybe also Belgrano) synchronize their observations with CALIOP overpasses, and this makes the comparison more reliable. I would suggest that the authors comment on the opportunity to perform synchronized measurements with CALIOP overpasses. The synchronized measurements do not improve the occurrence statistics necessarily, but they make comparison with CALIOP more reliable.

Please see our reply to your other comments regarding the possible measurement protocols at ground stations. We have now clarified that our data set corresponds to a CALIPSO-synchronous measurement protocol at ground stations. We have also dropped the misleading reference to continuously and manually operated instruments at ground and replaced the corresponding statements with more accurate ones.

Snels, ACPD, 2020: Snels, M., Colao, 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, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-972, in review, 2020.

Thank you. The paper has been added to the list of references.

Other comments:

Abstract. Line 8. What do the authors mean by representativeness ? Is it wrt to the CALIOP observations in a lat-lon box or wrt to the overall occurrence statistics in the Northern or Southern Hemisphere ?

The term representativeness in our study refers to the statistics derived without and with the effect of tropospheric clouds and other measurement-inhibiting factors that can affect the findings of ground-based lidar instruments. The statement was changed to: "*CALIPSO observations during the boreal winters from December 2006 to February 2018 and the austral winters 2012 and 2015 are used to assess the effect of tropospheric cloudiness and other measurement-inhibiting factors on the representativeness of ground-based PSC observations with lidar in the Arctic and Antarctic, respectively.*"

Line 12. These findings are rarely in agreement with polar-wide results..... Why would one expect an agreement with polar-wide results? Each location is different. It would be more interesting to have an agreement with a "box-region" observed by CALIOP

One might not expect an agreement with polar-wide results at each site but it is reasonable to assume that some sites are more representative of the larger-scale conditions than others. The agreement in a box region is basically what we do for the selection ground sites.

Line 15. Concordia is already a NDACC lidar observatory since 2014. Data are available on the NDACC web-site.

The Referee is correct. The statement has been changed to "and Mawson, Troll, and Vostok in the Antarctic". We rate Mawson over Jang Bogo due to the proximity of the latter to McMurdo.

Line 33 "calculations with" should read "calculations considering..."

Changed to: "light-scattering calculations that consider spherical and non-spherical particle shapes"

Line 43: representativeness see comment on line 8

Revised as in the reply to comment regarding line 8.

Line 47 : I would prefer "ground stations" instead of "ground sites", "site" already implies "ground"

Ground site has been changed to ground station throughout the text. However, we still use the term site when referring to locations.

Lines 81-83, This line is not very clear for readers that are not familiar with CALIOP data and should be written in a more "reader friendly" way. The 4 digits in the height are not significant and mentioning the bin number is irrelevant.

Thank you for this comment. The section was revised to: "*Because of CALIPSO's top-down viewing geometry, profile start with the uppermost height bin down (bin 1) to the lowermost height bin (bin 583). Profiles in the PSC mask v2 product extend down to 8.2 km. They can therefore contain contributions of upper-tropospheric cirrus, as visualised in Figures 13 and 20 of Pitts et al. (2018). To exclude the contribution of such cirrus clouds from our analysis, only height bins above 14.9 km (smaller than bin 85) and 13.1 km (smaller than bin 96) are considered to represent Arctic and Antarctic PSC, respectively.*"

Line 92 ..if this type....

The statement has been clarified to: "*Profiles are referred to as containing a certain PSC type (i.e. STS, NAT mixture, NAT enhanced, ICE or wave ICE), if this type was identified in at least one of the PSC height bins.*"

Line 101 . the 2x2 degrees boxes correspond with 220 x 76 km at 70 degrees of latitude and 220x38 km at 80 degrees latitude. This implies that the box dimensions change with the locations. Does this create a bias on the statistics ?

We don't think that this has much of an effect on the statistics as the smaller box sizes at higher latitudes are compensated for by the higher CALIPSO return rate at higher latitudes.

Line 108: I would add (iii) ground-based observations synchronized with CALIPSO overpasses.

Please see our earlier replies regarding the reference to ground-based measurements and the revisions of Section 2.4.

Line 108-113. The authors want to estimate potential biases due to the mode of operation of the ground based stations. The answer is apparently in the small numbers in Figure 7. To my opinion these numbers do not address adequately the question they posed in the introduction, since the difficulties encountered while recording ground-based measurements cannot be simply translated in doing random measurements. (implicating that non-random measurements would give different results..). "(ii) a manually operated system for which one third of the cases of the ground-based view was randomly selected." What does this mean and how it works? In most cases the number in the third column is about 1/3 of the second column, except for Tiksi. Why is that? What is the rationale between taking a random 1/3 or just divide by three ?

We are sorry about the confusion. We have revised Section 2.4 and hope that it is now more comprehensible. We do refer the bias to the available data coverage but erroneously assumed this could be synonymous to certain modes of operation. We actually only include what would be CALIPSO-synchronous measurements when we consider a CALIPSO lidar profile from the spaceborne and ground-based perspective. We then screen the data set to find those profiles in which tropospheric cloudiness was unlikely to have attenuated the lidar beam if it was coming from ground (no or only transparent clouds). This corresponds to the optimum data yield for ground-based measurements. However, we know that there is a wide range of factors that reduce the amount of collected data from the optimum data yield. We estimated that even under the worst of circumstances, a ground-based instrument should not provide less than one third of the maximum possible measurements. To get to this sub-set of observations, we randomly selected one third of those CALIPSO profiles that represent what would be observable from ground, i.e. the optimum yield. The statistics were derive subsequently from that subset of profiles. The numbers in Figure 7 refer to the number of PSC height bins in the corresponding category. Because the amount of PSC height bins can vary from profile to profile, the scaling is only about one third and not exactly one third.

Line 201. It is not clear what the 1:1 line means, and also the other grey lines like 1.0:1.6 are not clear. The authors write "the grey lines mark the ratios....." But which ratios ?

We agree that the grey lines in Figure 8 were confusing. We have removed all but one and revised the figure caption to: "*The grey line marks a scale PSC coverage defined as (10000 - x)/10000. Stations to the right of this line show a combination of tropospheric cloudiness and PSC coverage that indicates favourable conditions for ground-based lidar measurements*."

We have also revised the discussion of Figure 8 accordingly.

Line 202 add Concordia

done

Line229 understanding of processes.

of has been added

Figure 4 shows the occurrence rate of the different PSC classes as seen by CALIOP, by the ground-based lidar (continuously operating) in clear sky conditions and for manually operated ground-based stations. This figure is not clear for what concerns the small numbers written in the coloured columns. It would be better to have a Table with these numbers. Then the number of continuously operating lidars is very small.

We state in the figure caption that the numbers refer to the total amount of considered PSC height bins per configuration. The purpose of these numbers is to give an idea about the amount of data

that went into the respective bars. We have increase the size of the figure to improve readability. Please see our previous replies clarifying what was meant with the reference to continuously and manually operated ground-based instruments.

Figure 2 . The longitudes in fig b are wrong! / Figure 5 the longitudes are wrong in fig a / Figure 6 the longitudes are wrong

Thank you for spotting this mistake. It has been corrected.

Table 1. mark Concordia with existing datasets (see NDACC) The authors might indicate in Table 1 (or in a new Table limited to PSC observing stations) which lidars are continuously operated, which are randomly operated (whenever it suits the operator) and which are synchronized with CALIOP overpasses.

A corresponding marker has been added to Concordia in the table. We have also changed the marker style of Concordia in Figure 1b from open blue circle to filled magenta circle and in Figure 8 from open to filled circle to denote that it is a research station with published PSC measurements.