Interactive comment on “Depolarization ratio of Polar Stratospheric Clouds in coastal Antarctica: profiling comparison analysis between a ground-based Micro Pulse Lidar and the space-borne CALIOP” by C. Córdoba-Jabonero et al.

Anonymous Referee #3

Received and published: 8 January 2013

This paper compares depolarization measurements from a ground-based micro-pulse lidar (MPL) system with those from the CALIPSO spaceborne lidar system. The apparent goal of the study is to examine the performance of the improved ground-based system which now includes a built-in depolarization module and assess its capability for polar stratospheric cloud detection and classification. The comparisons of the MPL and CALIPSO depolarization measurements are based on two statistics: correlation coefficient and bias. The paper is well-written and of interest to the community. However, I have one serious concern with the analysis and some more minor issues that should be addressed before it is accepted for publication.

General Comments:

The conclusions indicate that although there is reasonable agreement between the datasets, there is a systematic bias between MPL and CALIOP with CALIOP depolarization being systematically higher. However, I think there may be a serious flaw in your analyses which may actually be causing at least part of this bias. You state that negative values of volume depolarization are disregarded in both datasets. I assume this means that negative values are thrown out when the CALIOP data is averaged. This is not the proper way to handle the CALIOP data. For low signal-to-noise systems like CALIOP, measurement noise can naturally lead to negative values for backscatter and hence, depolarization. These points should be included in scientific study. Any analysis that involves taking some form of average will exhibit a high bias if the negative points are excluded. For instance, the molecular depolarization for the CALIOP system is approximately 0.00366. If one examined a large amount of CALIOP depolarization measurements from molecular (cloud free) scenes only, the mean value of the depolarization would be near the value of 0.00366, but there will be a fairly broad distribution of points of which approximately half will consist of negative values. If these negative values are disregarded, the calculated mean molecular depolarization would clearly be biased high. So if the authors are disregarding negative values of depolarization when averages are being calculated, then they should redo their analyses to include the negative CALIOP values in the averages. This may reduce much of the observed bias.

A minor point of concern is the discussion of the various PSC types and their role in ozone depletion, both in the abstract and in the introduction. First of all, PSCs most likely occur in one of three particle compositions: super-cooled ternary (H2SO4, HNO3, H2O) solution (so-called STS or Type 1b), nitric acid trihydrate (NAT or Type...
1a), and ice (Type II). The authors never mention the liquid clouds in any of the relevant sections. In addition, the authors indicate in the abstract that ice clouds are the most important type of PSC for ozone depletion. Actually, the liquid STS PSCs probably are the most important for chlorine activation and the NAT PSCs are most important for denitrification. It probably would benefit the authors to read the nice PSC review paper by Lowe and MacKenzie (2008) and then rewrite these sections.

My other concerns with the paper are less serious and are listed below.

Minor Comments:

P.8054, L.23-27: I’m not sure if you’re implying here that a system with depolarization measurements alone would be sufficient for PSC detection. But just to clarify, as mentioned above, there is an important class of PSCs that consist of spherical liquid particles (STS) which will not produce any enhancement in depolarization. Although the depolarization measurement is important for separating spherical particles from non-spherical particles (i.e., STS from NAT or ice), it cannot be the only measurement used for PSC detection.

P.8057, L.18: CALIOP is not the first spaceborne lidar system. There was the Lidar In-space Technology Experiment (LITE) on the space shuttle in 1994 and the Geosciences Laser Altimeter System (GLAS) on ICESat that operated from 2003 to about 2009.

P.8058, L.4: How is this vertical averaging performed? Do you apply a running average over each 7 adjacent points in the vertical? What happens when your averaging window includes points with different vertical resolutions (i.e. across the 8.2 km or 20.2 km levels where the CALIOP vertical resolution changes)? What is the final resolution then of the CALIOP data after this vertical averaging? Are negative values disregarded when calculating these averages?

P.8058, L.1-4: Just a comment about the averaging scales of the two datasets in general. The ground-based lidar data is utilized as 1-hour averages with 75-m vertical resolution. The CALIOP data are 5-km horizontal x 7-point (???-m) vertical averages. Is the horizontal scale of the CALIOP data consistent with the hour-long averaged ground-base data? How does depolarization change on the time scale of one hour? How much cloud would have passed over the site during an hour (what are the typical wind speeds in the lower stratosphere over the station)? I would think it would correspond to much larger scales that the 5-km of CALIOP data. Maybe it is worthwhile to consider averaging the CALIOP data to appropriate scales to match those of the ground-based system.

P.8062, L.10-11: Again, you apply additional averaging to the ground-based data to improve SNR but not the CALIOP data. CALIOP likely has significantly poorer SNR than the ground-based system. How would the comparisons look if you didn’t apply the additional averaging to the ground-based data?

P.8062, L. 25-26: As noted above, disregarding negative values when calculating averages is incorrect. So if you do not include these negative values when calculating the 0.5-km layer averages, the averages are likely biased high.

P.8064, L. 1-30: It is not obvious to me what the correlation coefficient is telling you about the quality of the agreement between the datasets, especially how the agreement changes with altitude. The correlation coefficient will be dominated by the largest values in the profiles, i.e. the cloud layers. So I suppose the correlation coefficient is providing some measure of how well the profile shapes agree - is this correct? Could you provide more detail in the discussion to more clearly indicate what this analysis says about the quality of the agreement? A simple plot of mean/median differences and RMS differences as a function of altitude would be useful and maybe easier to interpret!

P.8065, L. 2-4: I find it surprising that the results have no dependence on distance between the station and the CALIPSO ground track, although in general a separation
of 55 km is still a very close coincidence. Is there the possibility to look at CALIPSO tracks that are even further than 55 km from the ground site? This may provide some insight to what spatial scales are important or if there is a problem with the analyses. I would expect as you increased the separation distance beyond 55 km, you would begin to see degradation in the quality of the agreement. Could you also use the other 56 coincidences where the ground-based measurements weren’t simultaneous to see how time differences impact the comparisons?

P8065, L. 14-15: I don’t understand how you can simply ignore differences that are larger than 50%—is there some justification for this?

P8068: I find Figure 5 difficult to read—would be useful to make it larger.

P8069, L. 1-27: It seems that the main conclusions are that there is ‘good’ correlation between the two depolarization datasets and ‘relatively good agreement’, but the MPL is biased low relative to CALIOP. I was hoping to see something more quantitative about the quality of the agreement. Can you provide any more quantitative information here? It seems a simple statistical analysis of the 48 cases and examination of the median and/or mean differences (as a function of altitude) and the standard errors would be useful and maybe complement the CC and bias analysis.

References
