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***

Authors strongly appreciate the helpful comments of the referee #3 with valuable suggestions and useful remarks, patently improving this work.

We will try to answer and explain, in any case, all the questions. In particular, the previous lidar datasets and new data have been processed and re-analysed in order to respond the questions of all the referees. Indeed, negative depolarization values as proposed by the referee #3, within given restriction limits (see response to all his/her comments, next), as well as new data profiles corresponding to higher CALIPSO overpass distances from Belgrano II station than those already considered, as requested by both the referees #1 and #3, have been also processed and included in the comparison analyses between lidar datasets. In particular, several changes (three Figures have been added: the new Figures 2, 3 and 6; old Figures 2 and 3 have been modified; old Figure 4 replaced; old Figure 5 splitted into 4 separated figures with other particular cases selected: the new Figures 8, 9, 10 and 11; Table 2 removed, and Tables 3 and 4 have been modified and joined together in one Table, the new Table 2) have been included in the manuscript as a consequence of responding to all the referees' comments. Then, Sections, Figures and Tables have been renumbered. As a result the text has been accordingly modified to contain these changes, including new required calculations, analyses and results. In general, a revised manuscript containing all the necessary modifications is also available.

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:

GC 1)
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.

<u>Authors</u>: We appreciate all the comments of the referee, especially his/her important remarks regarding CALIPSO data processing. It is true that negative values cannot be disregarded when SNR is low as for both CALIOP and MPL-4. But after a first look to the CALIPSO depolarization data, we observed high negative values inside the CALIPSO dataset, even after smoothing procedures, and mainly at higher altitudes. That's why authors decided to disregard negative values in order to avoid the 'contamination' of data with such 'unrealistic' values, at least at a first place. However, considering the referee's concern on this subject, we have performed a re-analysis of all our comparison study. However, we have considered a restriction to the negative values to be used for that comparison analysis following Pitts et al. (2009). In this work, aerosol depolarization ratio (δ_a) data between -0.1 and 0.8 were shown, falling about 14% overall of CALIOP data outside these limits. Taking into account that we present the volume linear depolarization ratio (δ^V) instead of δ_a, we have calculated the corresponding limits by using the formulation reported by Cairo et al. (1999), linking both these magnitudes. Hence, δ^V can be expressed as a function of the molecular depolarization ratio (δ_{mol}), the backscattering ratio (R) and δ_a, as follows:

$$\delta^V = \frac{\delta_a \times (R \times \delta_{mol} + R - 1) + \delta_{mol}}{\delta_a + (R-1) \times \delta_{mol} + R}$$

where R values are varied between 1 and 30, δ_a between -0.1 and 0.8, and δ_{mol} takes these two values: 0.00366 (Cabannes scattering) and 0.0144 (total Rayleigh). Among all the possible combinations, the minimal (-0.1) and maximal (0.8) values are computed. Hence, these limits will act as a conservative restriction range for the overall of δ^V values in both lidar datasets that is the same limits presented for δ_a in Pitts et al. (2009). In fact, about 14 ± 11% and 5 ± 3% overall of CALIOP and MPL-4 data, respectively, falling outside of these limits, have been disregarded in our re-processing of lidar depolarization data.

All this information appears now in the revised version of the manuscript, introduced within the Sect. 2.2.2.

<u>References</u>:

Cairo, F., Di Donfrancesco, G., Adriani, A., Pulvirenti, L., and Fierli, F.: Comparison of various depolarization parameters measured by lidar, Appl. Opt., 38, 4425-4432, 1999.

Pitts, M.C., Poole, L.R., and Thomason, L.W.: CALIPSO polar stratospheric cloud observations: second-generation detection algorithm and composition discrimination, Atmos. Chem. Phys., 9, 7577-7589, 2009.

GC 2)
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 (H_2SO_4, HNO_3, H_2O) 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.

<u>Authors</u>: The scope of this work is not the PSC-type discrimination (it will be done in the future), but the performance of the depolarization measurement capabilities of such a kind of ground-based Micro Pulse Lidars (MPL) for PSC observations. Previously, a comparison against a ground-based lidar devoted to long-term PSC measurements was already performed in the Arctic with the same lidar system (Córdoba-Jabonero et al., 2009), but without a final conclusion on such MPL depolarization capabilities. Therefore, a new MPL depolarization inter-comparison with a reliable lidar located not far from the Belgrano II station was necessary. At polar latitudes, such a lidar couldn't be other than the space-borne CALIOP. Therefore, the expecting comparison between both lidar datasets has been presented and analyzed in this work. Authors didn't want to deeply address the PSC classification in this paper, also because backscattering ratio (R) data are not shown regarding the scope of the paper as said before. In fact, a future work is to be performed following the current one addressing this question. However, it is true that more information should be provided on PSCs, at least, in the abstract and introduction. For instance, in the reference mentioned by the referee (Lowe and MacKenzie, 2008), which has been also included in the manuscript's reference list, a more extensive information can be found on this subject.

In this sense, authors have rewritten both the abstract and introduction including more information on PSC classification, their types and compositions, and their corresponding role in ozone depletion. In particular, both PSC-I (subtype Ia corresponding to solid particle clouds and subtype Ib to liquid clouds) and PSC-II (ice clouds), as well as their mixtures, have been commented.

<u>References</u>:
Córdoba-Jabonero, C., Gil, M., Yela, M., Maturilli, M., and Neuber, R.: Polar Stratospheric Cloud observations in the 2006/07 Arctic winter by using an improved Micro Pulse Lidar, J. Atmos. Ocean. Technol., 26, 2136-2148, 2009.

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

Minor Comments:

MC 1)

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.

Authors: Lidar systems insensitive to polarization are sufficient for PSC detection (for instance, see the work on five-years of MPL measurements on PSC detection at South Pole station by Campbell et al. (2008), and references therein); however the PSC-type discrimination additionally requests the use of depolarization capabilities in order to analyze both backscattering ratio (R) and depolarization ratio (δ^V). Authors consider indeed this paragraph was ambiguously written in the previous version of the manuscript. As referee #3 claims now, it has been reworded as follows:

"Lidar measurements have been widely used for PSC classification on the basis of two lidar variables: the backscattering ratio (total backscatter-to-molecular coefficient ratio, R) and the volume linear depolarization ratio (δ^V). Indeed, due to the fact that non-spherical particles change the polarization state of the incident light, unlike spherical particles, both PSC-I (subtype Ia corresponding to solid particle NAT clouds and subtype Ib to liquid STS clouds) and PSC-II (ice clouds) as well as their mixtures, can be detected and identified by using lidar systems with depolarization measurement capabilities."

References:

Campbell, J.R., and Sassen, K.: Polar Stratospheric Clouds at the South Pole from Five Years of Continuous Lidar Data: Macrophysical, Optical and Thermodynamic Properties, J. Geophys. Res., 113, D20204, doi:10.1029/2007JD009680, 2008.

MC 2)

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.

Authors: Completely in agreement with this comment. Then, the manuscript has been accordingly modified.

MC 3)

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?

Authors: The CALIOP level 1B data used in this work present a vertical resolution of 30 m, 60 m and 180 m at altitudes lower than 8.2 m, between 8.2 and 20.2 km and 20.2 *and 30 km, respectively. This change in averaging scales is performed in order to reduce the noise level of the profiles. As mentioned in the manuscript, a vertical 7-point adjacent averaging is applied in order to reduce the noise level of the profiles (a horizontal averaging over 5 km CALIPSO ground-track is also performed for that purpose). I.e., a data smoothing is achieved by using a sliding window of 7 points for averaging through the entire profile. As a result of this procedure, the final resolution of the CALIOP data keeps unchanged.*

That change in averaging scales across both the 8.2-km height and 20.2-km height boundaries would produce unpredictable results depending on the particular atmospheric situation. However, in general, this smoothing process doesn't significantly impact our following '0.5-km averaged' δ^V profiles from CALIOP because the 0.5-km averaged height-range, where an inhomogeneous smoothing is applied (in terms of vertical resolution but maintaining the number of points), cover the small atmospheric regions between 8.08 and 8.56 km and 20.08 and 20.56 km in each boundary case. Note that negative δ^V values are not disregarded during the smoothing process of CALIOP data profiles (see Sect. 2.1.2). Only δ^V values falling outside the (-0.1, 0.8) interval are ignored for the following 0.5-km vertical averaging of both lidar profiles (see response to the referee #2's Comment 2).

The manuscript has been accordingly modified by including a more complete explanation of the smoothing and averaging procedures applied.

MC 4)
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.

<u>Authors</u>: *First, we would like clarify that the hourly-averaging performed to the MPL-4 δ^V profiles is based on a SNR enhancement for these lidar measurements, mainly at altitudes > 20 km height, where the noise considerably increases. Temporal fluctuations of these δ^V values during that hour are considered included within the error analysis for that parameter. However, in relation with this referee's comment, and for a more clarification, an analysis of the time averaging and temporal fluctuations of the MPL-4 depolarization profiles has been included in the new revised manuscript, rewriting the Sect. 2.2.1. The following text has been introduced:*

"*Regarding time averaging procedures applied to the MPL-4 measurements, hourly-averaged MPL-4 δ^V profiles, as obtained from those 30-min averaged p- ($P^{\parallel}(z)$) and s- signal ($P^{\perp}(z)$) profiles in one hour (see Sect. 2.1.1 for details), are those used in the comparison with CALIOP data instead of instantaneous 1-min profiles ($\delta^V{}_{1-\min}$). As aforementioned (see Sect. 2.1.1), this improves the SNR of the lidar measurements at Belgrano II station. Indeed, the level of noise decreases as the time averaging* *increases, as shown in Figure 2 (for instance, data on 1 July 2009). δ^V variations depending on the time averaging (5-, 10-, 15- and 30-min averaged profiles are shown in Figure 2) reveal that the vertical δ^V structure presents a clearly enhanced SNR when the time averaging is higher than 15 min. In particular, an additional PSC feature at around 25 km height can be identified with enough SNR only for 15-min and 30-min averaging of the MPL-4 data (see Figure 2).*

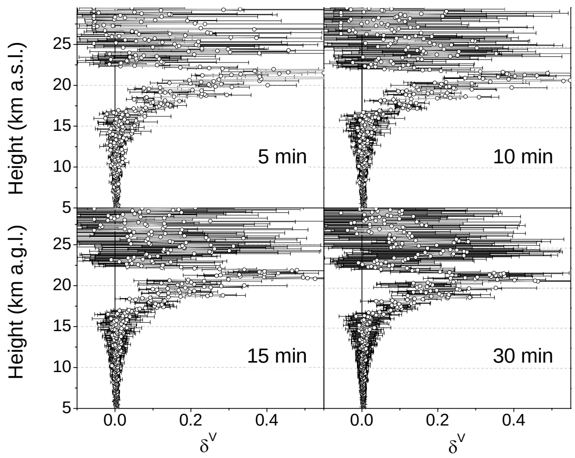

Figure 2. MPL-4 δ^V profiles (grey-lined open circles, in black background of their SD values) on 1 July 2009 depending on the time averaging (see legend inside each panel, from up to down and right to left): 5-, 10-, 15-min and 30-m averaged profiles.

Moreover, MPL-4 δ^V fluctuations along that hour are also studied by examining the differences between instantaneous 1-min ($\delta^V_{1-\mathrm{min}}$) and hourly-averaged ($\delta^V$) profiles within the same hour. Mean differences and their RMS values are shown in Figure 3 (for instance, data on 1 July 2009). A height-averaged value of -0.005 ± 0.013 is obtained for these mean differences, and their RMS values show that temporal δ^V fluctuations are lower than 0.05, 0.1 and 0.25 up to altitudes of 18 km, 23 km and 30 km height, respectively.

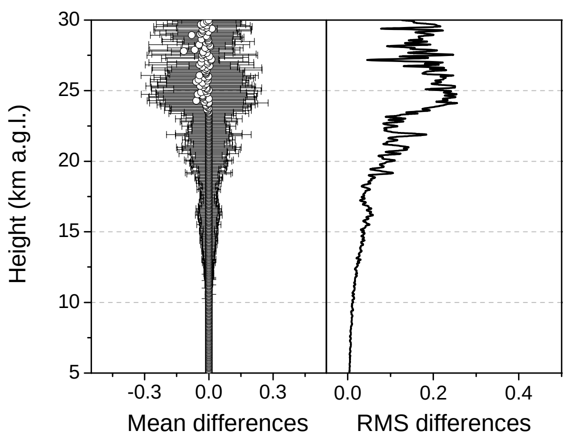

Figure 3. MPL-4 mean ($\delta^V{}_{1-\min} - \delta^V$) differences along an hour and their RMS values (for instance, data on 1 July 2009).

In summary, a high SNR is achieved for a time averaging of 30 minutes applied to MPL-4 p- and s- signal profiles (i.e., hourly-averaged δ^V profiles) in our comparison analysis, and temporal δ^V fluctuations per hour are lower than 0.1 up to altitudes where PSC features more frequently appear."

The manuscript has been accordingly modified in order to include this kind of information (these new Figures 2 and 3 have been added), thus providing a more complete clarification. Evidently, all these modifications will improve the manuscript.

Moreover, regarding the referee's question: "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)?", *authors have calculated the mean monthly-averaged wind velocities (m/s) from 2009-2011 winter periods (data from local radiosoundings launched in Belgrano II station, see Figure attached below). As you can observe, wind velocities can vary month-to-month. At low stratospheric altitudes, between 10 and 25 km height, the minimal-maximal values for each month are: 11-31 m/s, 18.5-29 m/s, 15-22 m/s and 17.5-24 m/s, respectively, from June to September, and the largest variation (20 m/s) is found for June in this height-range. However, in order to properly address this issue an analysis of other suitable variables would be required, as for instance the wind directions in addition to their speeds, and furthermore backtrajectory computations would be valuable for that study. Nevertheless, all these additional analyses are out of the scope of this work.*

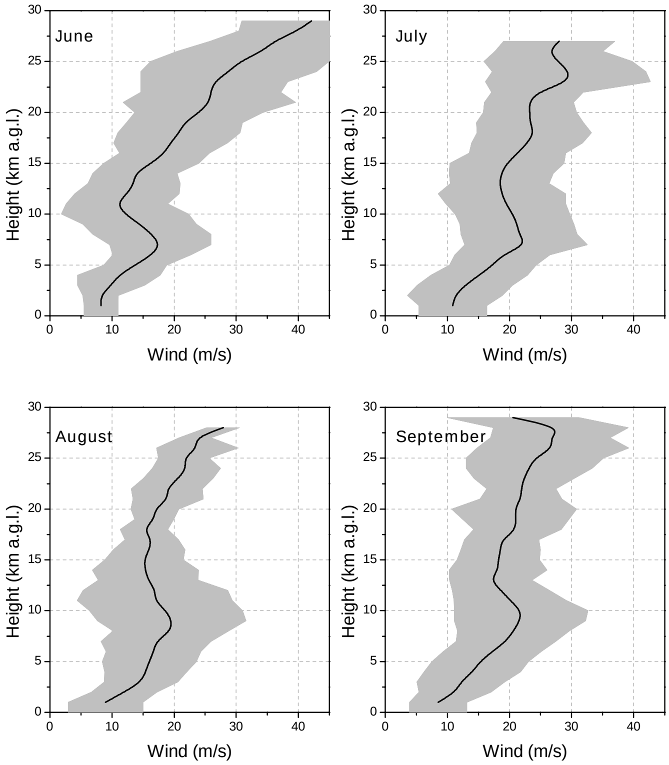

Figure. Mean monthly-averaged wind velocities (black line) from the 2009-2011 winter periods (SD is shown by the grey-shaded band). Data are from local radiosoundings launched in Belgrano II station.

MC 5)
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?

<u>Authors</u>: We disagree with the reviewer #3's comment regarding the application of procedures to improve the SNR for one system but not for the other one. Both lidar datasets have been averaged in time to improve the SNR. On the one hand, ground-based data have been averaged over 1 hour (i.e., 30 of those 1-min instantaneous profiles corresponding to each polarization-sensitive channel). On the other hand, CALIOP data have also been averaged over 5 km, corresponding to 15 individual profiles registered during the time interval needed to overpass those 5 km long (then data are smoothed as explained before; see response to the referee's Minor Comment 3).

MC 6)

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.

<u>Authors</u>: As aforementioned, re-calculations have been performed including negative values, but within a restriction limit interval of (-0.1, 0.8) (see response to the referee #3's General Comment 1). New results have been included in the manuscript, accordingly modifying the text.

MC 7)
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!

<u>Authors</u>: The correlation coefficient (CC) of a statistical model (i.e., a linear regression) describes how well a set of observations are fitted each other. In our case, the MPL-4 vs. CALIOP comparison, the CC is a measure of the goodness of that assumed linear model, quantifying the discrepancies between the values reported by both the lidar systems. Therefore, the CC of a linear regression will indicate the goodness of the proposed linear fit, and in particular CCs higher than a significant value (i.e., 0.5 was chosen in our work) will indicate a good agreement in terms of profile shapes, i.e., a measure of good relationship between both cloud structures as a function of altitude.

Moreover, following the referee's suggestions, mean differences between both lidar datasets and their root mean square (RMS) values as a function of altitude have been calculated and the obtained results included in the text, complementing thus the previous both comparison analyses. All this information appears now in the revised version of the manuscript as a new subsection within Section 3.2 ("3.2.2 Comparison analysis II: Mean differences (Δ) and their root mean square (RMS) values").

MC 8)
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?

<u>Authors</u>: In order to answer this referee's comment, two questions would be to be addressed: 1) the examination of the results when new data for distances higher than 55 km are included in the comparison analysis; and 2) the effect of no-simultaneous measurements to see how temporal differences impact the comparisons. The latter

question has been partially addressed previously (see response to the referee's Minor Comments 4 and 5); in addition, authors consider that no-simultaneous measurements of two hours before or later the CALIPSO overpass are out the scope of this work, based on requirements of any intercomparison: coincidence in time and space. Therefore, analysis comparison between both lidar datasets has been restricted, at least, in time in this work, and the comparison analysis has just been performed for simultaneous measurements carried out at time scales lower than two hours around the CALIPSO overpass. However, regarding coincidence in space, new data have been analyzed to include the influence of spatial scales when data from rather high CALIPSO overpass distances are included in that lidar profile comparison.

The manuscript has been accordingly modified in order to include this kind of information (old Figures 3 and 4 have been modified), expecting that it will provide a more complete clarification. Evidently, all these changes will improve the manuscript. In particular, Section 3.1 has been rewritten as follows in the revised version of the manuscript:

"PSC observations have been performed at the Belgrano II station since 2009 to the present. MPL-4 measurements for the 2009-2011 Antarctic winters, from May to September, are used for this study. Lidar datasets are compared under the following conditions: coincident profiles in time, with simultaneous measurements carried out at time scales lower than two hours around the CALIPSO overpass, and in space, with CALIPSO ground-track separations from the station closer in 55 km distance. During those winter times, a total of 189 CALIPSO overpasses nearby the Belgrano II station were carried out within fewer 55 km distance. Among them, 104 overpasses are coincident events with MPL-4 measurements reporting PSC detection and 48 of them are simultaneously available for comparison. Moreover, 35 more lidar profiles are analyzed in order to examine the influence on spatial scales when data from rather large CALIPSO overpass distances are included in that lidar comparison. Those additional 35 profiles correspond to separations between 70 and 100 km. In general, four predominant distance-ranges are observed: 0-10 km, 20-30 km, 45-55 km and 70-90 km. All these PSC cases, listed by the CALIPSO ground-track distance from the Belgrano II station, are shown in Table 2.
A height-interval from 5 to 30 km is selected for the comparison between lidar profiles. A delineating altitude of 10 km has been conservatively established as the lower limit for the unambiguous presence of PSCs, distinguishing them from other upper tropospheric clouds (mainly Cirrus clouds). The lower limit of 10 km chosen in this work for PSC detection is based on the fact that the tropopause is not clearly delineated by the temperature profile during wintertime in deep Antarctica (Rubin, 1953). Indeed, a traditional tropopause height, denoted by rapidly increasing static stability above it, can be approximated from December through March in Belgrano II station sounding data around 9 km. During winter months, however, temperatures decrease with height to nearly 23 km. Dynamic coupling between the troposphere and stratosphere is more likely in such conditions. The region from 8 to 10 km is considered a transitional zone, where cloud type cannot be established with any certainty. Although our study is restricted to PSC formation altitudes, the 5-10 km height interval is also considered for contrast as a PSC-free region. Heights above 30 km are disregarded due to a decreasing of the MPL-4 SNR. Finally, along the overall height-interval selected, every 0.5-km layer is averaged for comparing the MPL-4 and CALIOP datasets. Note that negative δ^V *values are not disregarded during the smoothing process of CALIOP data profiles (see Sects. 2.1 and 2.2). Only* δ^V *values*

falling outside the (-0.1, 0.8) interval are ignored when this 0.5-km vertical averaging is applied to both lidar profiles."

MC 9)
P.8065, L. 14-15: I don't understand how you can simply ignore differences that are larger than 50%- is there some justification for this?

<u>Authors</u>: *The application of a determined constraint conditions to the percentage differences ($BIAS$), which the latter analysis discussion is based on, between both lidar datasets is justified by the large data dispersion found within the overall $BIAS$ values, being higher than ±50%. This is in relation with the fact that $BIAS$ is rather large when relatively low depolarization values are present within a given CALIOP (δ^{CAL}) profile (as, for instance, those provided at PSC-free altitudes). And as a consequence, the $BIAS$ considerably increases (see Eq. (7)), exceeding the percentage values of that supposed constraint condition.*

The manuscript has been accordingly modified in order to include this kind of information, providing thus a more complete explanation. In particular, the following paragraph has been included in the Section 3.2.3 of the revised version of the manuscript, replacing the previous one:

"Due to the large $BIAS$ data dispersion obtained in general, a determined constraint condition is applied to the $BIAS$ profiles: only $BIAS$ values within a given interval (-50% < $BIAS$ < +50%) are regarded. This constrained selection is done to evaluate $BIAS$ values with a given realistic significance, ignoring values rather higher than ±50%. From a statistical point of view, a height-averaged $BIAS$, $BIAS_z$, is calculated from the 'constrained' profiles, also considering the number of data points in each profile fulfilling that condition. 'Constrained' $BIAS_z$ values for all the cases are shown in Table 2 together with their standard deviation (SD), including the percentage of number of data points fulfilling that condition (N^{BIAS}, in %) for each case."

MC 10)
P.8068: I find Figure 5 difficult to read- would be useful to make it larger.

<u>Authors</u>: *Old Figure 5 (from top to bottom panels) has been replaced by four separated figures with other particular cases selected: the new Figures 8, 9, 10 and 11. As a consequence, all the Figures have been renumbered. Corresponding changes have been accordingly introduced in the text (Sect. 3.2.4).*

MC 11)
P.8069, 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.

<u>Authors</u>: Regarding all the referees' comments and their useful suggestions, new analyses have been performed, in particular including both negative values within a given restriction limits (see response to the referee's General Comment 1) and new data from distances higher than 55 km. In addition, mean differences together with their root mean square (RMS) values have been also calculated, complementing the results obtained by the two previously applied analysis approaches. Hence, the text (conclusions included), tables and figures have changed. All these modifications appear now in the revised version of the manuscript, improving patently this work.

In particular, conclusions about the degree of agreement between both lidar depolarization datasets have been established quantifying the comparison analysis by means of several approaches as the correlation between both vertical layering structures and both the mean and percentage differences (this latter procedure has still been included in our comparison analysis because this approach is usually used in profiling comparisons between CALIOP and other ground-based lidar systems). Regarding the results obtained now, they are discussed in more detail, thus providing more complete information. Therefore, the text of this section has been modified as follows:

"This study appears as a first application of the lidar depolarization technique to Antarctic PSC detection and identification by using an improved version (MPL-4) of the standard NASA/Micro Pulse Lidar. In particular, this work represents a significant advance on PSC-type discrimination studies by using MPL-4 δ^V data.

Calibration parameters for suitable MPL-4 δ^V retrievals have been calculated from MPL-4 measurements. These calibrated δ^V profiles have been compared with coincident CALIPSO data as a reference during 2009-2011 austral winters, from May to September periods, over Belgrano II station (Argentina, 77.9ºS 34.6ºW, 256 m a.s.l.). Coincident observations are referred to simultaneous measurements between both lidar systems within two hours around the closest CALIPSO overpass to the Belgrano II station. That is, four predominant distance-ranges between the CALIPSO ground-track and the station have been selected in order to examine the dependence of the degree of agreement between both lidar δ^V-profile datasets on that distance from the Belgrano II station. Three analysis procedures for δ^V-profile comparison between both lidar datasets have been presented.

Correlation analysis shows that a 59% out of all the comparison cases present CC values higher than 0.5, at least in the altitude-range from 5 to 20 km height. This frequency (number of cases with CC > 0.5) decreases down to 35% when the overall height interval of 5-30 km is considered for correlation fitting, indicating that discrepancies between CALIOP and MPL-4 in vertical layering structure are enhanced from 20 km up, likely due to a decrease of the SNR for both lidars systems at those altitudes. However, PSC formation occurs indeed more frequently at altitudes lower than those heights. Hence, a relatively good agreement is found between both ground-based MPL-4 and space-borne CALIOP profiles of the volume linear depolarization ratio δ^V for PSC events, once the MPL-4 depolarization calibration parameters are applied. Moreover, no large differences are unexpectedly observed in the correlation analysis between both vertical depolarization structures when the CALIPSO ground-track is just at a few kilometres from the station, showing a frequency of 75% with CC > 0.5, respect to rather large separations with a lower frequency (49%).

Regarding the differences between both lidar δ^V -profile datasets, two related variables are analysed: the mean differences, Δ, within a given CALIPSO distance-range; and the percentage differences, $BIAS$ (see Eq. (7)), since, despite its relation with the former Δ, $BIAS$ is a parameter also used in profiling comparisons between CALIOP and other ground-based lidar systems.

Slightly discrepancies are observed for the mean differences between those δ^V profiles depending on the CALIPSO separation. Mean differences Δ are mostly negatives in the overall height interval, with higher data dispersion at altitudes from 25 km height up. In addition, absolute Δ values are no higher than 0.1 in average, being lower than 0.05 at altitudes up to 25 km height, with RMS values no larger than 0.3 in overall for all the distance-ranges. The other comparison parameter, $BIAS$, seems to be a less robust indicator of the degree of agreement for lidar δ^V datasets, since rather less than a half of the cases fulfil the selected constraint condition (-50% < $BIAS$ < +50%), showing a large data dispersion with absolute percentage differences higher than 50%. This is related to the fact that $BIAS$ is rather large when relatively low δ^{CAL} values are present within a given CALIPSO profile (as, for instance, those are at PSC-free altitudes), and as a consequence, $BIAS$ considerably increases (see Eq. (7)), exceeding the percentage values of that supposed constraint condition. In addition, that previously observed predominance of negative values is also found, indicating a generalized δ^{MPL} underestimation with respect to CALIOP data. However, absolute differences between δ^{MPL} and δ^{CAL} are no higher than a 10 $\pm$ 11% in average as compared to CALIOP values. As expected, these results are in agreement with those obtained for the mean differences.

Moreover, the degree of agreement between both lidar δ^V datasets is moderately dependent on the CALIPSO ground-track overpass distance from the Belgrano II station, as shown by the results obtained in each one of the comparison analyses carried out: the vertical correlation (CC), and both the mean (Δ) and percentage ($BIAS$) differences. That is, no large discrepancies are found when CALIPSO ground-track distance is as close as < 10 km far as well as rather far (at 70-90 km) from the Belgrano II station.

Actually, these results indicate that MPL-4 depolarization observations would reflect relatively well the PSC field that CALIOP can detect at large distances from the ground-based station. As a consequence, PSC properties would be statistically similar in average over large volumes, and hence the present disagreement found between both the lidar δ^V datasets would be likely related to be dominated by small spatial PSC inhomogeneities along the CALIPSO separation from the station. This statement is based on the fact that Belgrano II is a station located well inside the polar vortex during almost all the wintertime period. Indeed, the Antarctic polar vortex is quite stable to allow determined thermodynamic conditions leading to a very low variability of the PSC field, and then of their properties.

Therefore, a further study, out of the scope of this work, on the correlation between the PSC features and the variability of both the polar potential vorticity and the stratospheric temperature fields would provide an understanding on what the observed discrepancies are really based on. In addition, a detailed going-on 3-year statistical analysis of PSC occurrence over Belgrano II station in terms of both the *backscattering, R, and volume linear depolarization, δ^V, ratios from MPL-4 measurements would complete these studies. This will include a PSC-type discrimination assessment over Belgrano II station, a station well inside the Antarctic polar vortex.*

Finally, it is worth mentioning these results are useful for PSC detection and classification in both Polar Regions by using this kind of micro pulse lidar that operates in full-time continuous mode, providing a more complete evolution of the PSC field on a daily basis."

References

Lowe, D. and MacKenzie, A. R.: Polar stratospheric cloud microphysics and chemistry, J. Atmos. Solar-Terr. Phys., 70, 13-40, 2008.