Answers to J. Joiner, A.Vasilkov, P.K. Barthia

This paper presents, to our knowledge, a unique retrieval using the O2-A band in which 2 pieces of information are retrieved: Geometrical cloud-top height and thickness. In addition, the geometrical cloud fraction is from a DLR algorithm based on analysis of the Polarization Measuring Device (PMD) data of GOME-1. Cloud optical thickness is then calculated using a wavelength of 758 nm outside the O2 A band. As this algorithm is unique, it is of great importance that the resulting retrieved cloud parameters be thoroughly validated.

General comments
In Sect. 2, there is discussion on the fact that when a cloud is modeled as a Lambertian diffuser, that a cloud height retrieval provides a value closer to the altitude of the middle of the cloud. This is certainly a true statement that has been recognized for some time as stated in the manuscript. However, the paper may leave a reader with the impression that cloud top height is the goal of these retrievals. The OMI science team has long recognized that a single piece of information retrieved using similar approaches with UV and visible measurements is not the cloud-top pressure. Note that for OMI two methods are used - rotational-Raman scattering in the UV and oxygen dimer absorption in the visible. We would like to bring to your attention several relevant papers on this topic that have not been referenced.

From the start of the mission, the OMI cloud pressure retrievals were not referred to as cloud top pressure, but rather as "effective cloud pressure". Since then, we have adopted a more descriptive term: the "optical centroid pressure" (or OCP), recognizing that what we retrieve is more similar to a reflectance-weighted pressure (please see papers by Vasilkov et al., 2008 and Sneep et al., 2008). Our latest paper (Joiner et al., 2011) discusses this in detail and provides a fast simulator that predicts the OCP given a profile of cloud optical extinction that can be vertically inhomogeneous or even absorbing. Vasilkov et al. (2004), Ziemke et al. (2009), Joiner et al. (2009), and Vasilkov et al. (2010) exploit cloud OCPs to provide accurate retrievals of total and tropospheric ozone, in particular over snow/ice and to derive ozone concentrations inside the tops of deep convective clouds (in combination with measurements from the Aura Microwave Limb Sounder). In these papers, it is shown that gaseous absorption takes place throughout the volume of a cloud (with many types of clouds being vertically inhomogeneous) and does not stop at the cloud top level. It is demonstrated that the cloud OCP concept (within the context of the mixed Lambertian model) correctly accounts for this absorption if the trace gas is well-mixed within the cloud volume. For convective clouds this is a reasonable assumption. The cloud OCP concept is therefore also appropriate for short-wave flux calculations (Joiner et al., 2009; Vasilkov et al., 2009) whereas the cloud-top pressure is more important for long-wave fluxes. If the trace gas is not well mixed inside the cloud(s), the vertical structure of the clouds becomes important and cannot be accounted for by a vertically homogeneous cloud model. In addition, Joiner et al. (2010) shows how cloud OCP may be combined with coincident cloud-top pressure from thermal IR measurements to detect multi-layer clouds. This approach was uniquely validated

by using nearly coincident CloudSat radar measurements along with OMI and MODIS data. The paper shows that in some areas, such as over the tropical Pacific, the fraction of cloudy pixels containing distinct multilayer clouds can frequently be 50% or more at OMI spatial resolution. This paper also shows that the frequency of multi-layered clouds in a pixel increases with pixel size. The occurrence of distinct multi-layer clouds should therefore be a significant issue for GOME-1 given its much larger pixel size.

This leads to a more general question about the SACURA approach - how does the algorithm behave when clouds are vertically inhomogeneous, not only in multi-layer clouds but also in deep convective and other types of clouds? After looking at a large number of CloudSat profiles, we find that the condition of vertical homogeneity is rarely met (see Ziemke et al., 2009 showing average cloud extinction profiles for tropical deep convective clouds that peak at different pressures depending in general on the total optical thickness). This issue has indeed been examined in previous papers by coauthors. However, a more extensive and detailed simulation would be beneficial.

(A.1) The retrieval procedure of the SACURA algorithm takes into account the vertical inhomogeneity of clouds by recursively finding the single scattering albedo profile $\omega_0(z)$ through the cloud, along the oxygen A-band. The optical thickness is calculated at $\lambda = 758$ nm, where ω_0 almost equals 1 at all heights. Equations and the error analysis can be found in Kokhanovsky and Rozanov [2004].

Addressing the issue of multi-layer clouds, we have run synthetic tests for a two-layer system with the lower water cloud of COT 10, CBH-CTH 3-4 km. In the first case, the upper water layer was fixed at heights 13-15 km; in the second, an ice cloud was simulated with a fractal crystal model of 50 μm side length, and placed at 13-15 km as well. The value has been chosen from CALIPSO dataset, as reported in Sassen et al. [2008], Fig.2. Solar zenith angle was set 30°, consistent with tropical latitudes. With increasing optical thickness of the upper layer, the curves show the cloud bottom (red curve) and top (blue curve) height retrieved values of the lower layer. The lower panel shows the total COT retrieved for both layers.
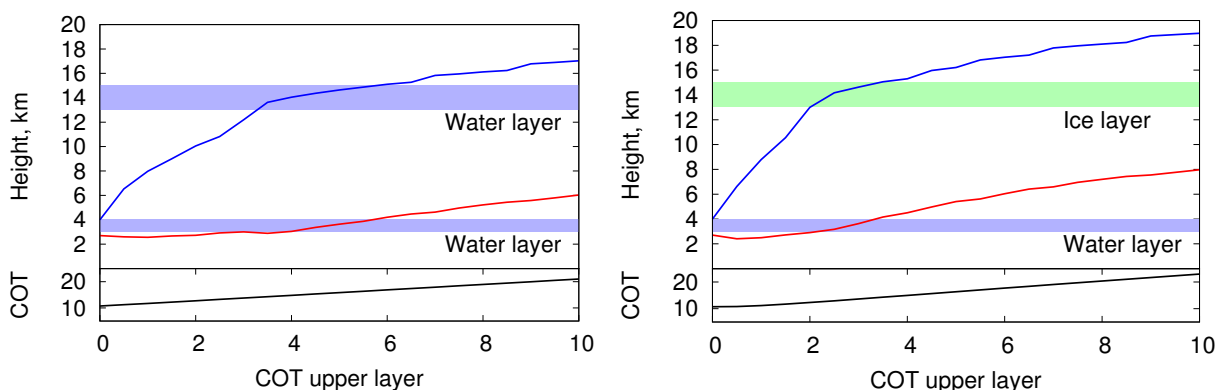
Looking at the retrieval flags, we notice that the operational limit of geometrical thickness (11 km) is met at $COT_{water} = 4$ and $COT_{ice} = 2$. Beyond that value, CTH and

2

CBH are constrained, and all successive retrievals are flagged 3. Given the occurrence of multilayered systems, we would then reject retrievals flagged 3, above a limit height of 5 km.

In Figure 1 we plotted the occurrences of the quality flags in function of CTH from the SNGome dataset, and we found that the situation reproduced in the above simulations contributes to the 2nd mode of the green curve. In fact, we are now able to apply a better selection on the data and to extract a more realistic set. Eventually the reanalysis of all plots exhibits more low level clouds.
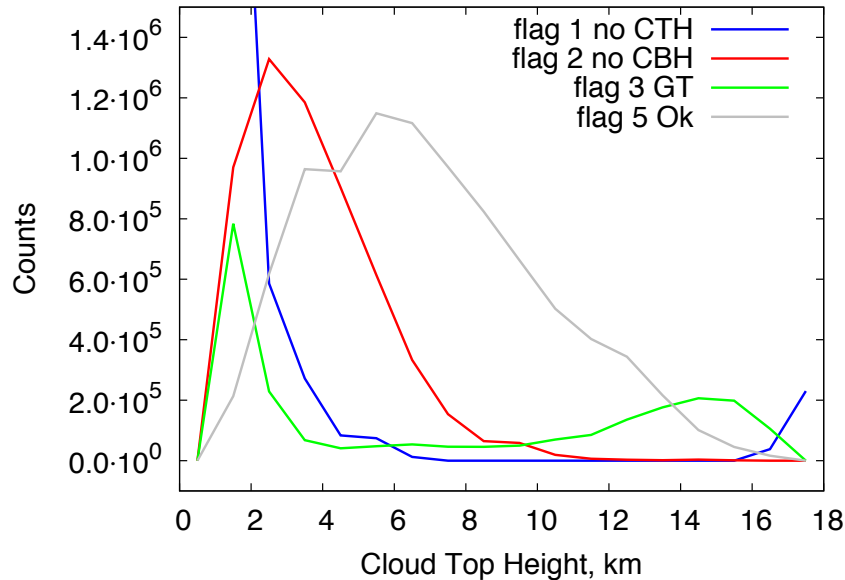

Figure 1: Quality flag counts versus CTH for the 7 years of SNGome dataset.

We have a set of a few thousand representative CloudSat extinction profiles from a single day that we would be happy to share with you. We used these profiles for our own simulation study in Joiner et al. (2011).

(A.2) It will be a valuable comparison and we are thankful to the referees for making the dataset available, and we will contact you as soon as possible.

One of the unique aspects of the SACURA algorithm is that it attempts to retrieve more than one piece of information about cloud vertical structure. Most of the paper is devoted to CTH, but the algorithm also provides an estimate of cloud base. This can be compared with cloud base from the ground-based measurements (radar, lidar, ceilometer) that should be quite accurate. Is the cloud base from your algorithm accurate or is it mainly error sink that allows for a better CTH estimate? For instance, what happens when clouds become optically thick in the middle and very little light penetrates to the cloud base?

(A.3) In Lelli et al. [2011] we have presented a sensitivity study with the following value ranges: COT in [5–50], CTH in [1–16] and CBH in [0–15]. A 1 km single-layer cloud was assumed, over a black (SA = 0) and a moderately bright surface (SA = 0.2). We report here for the sake of completeness the figures for CTH and CBH. In our model

Cloud Top Height absolute error, km

Cloud Bottom Height absolute error, km
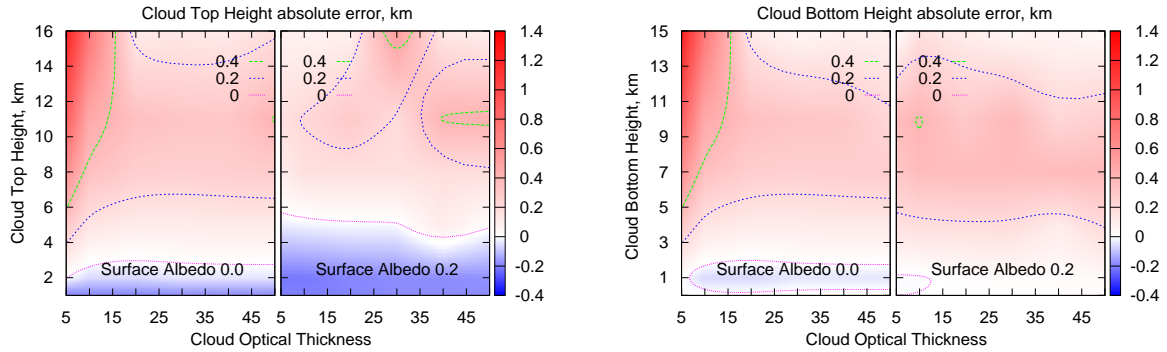
the cloud top and bottom heights are reciprocally linked by optical thickness, therefore the errors in the retrieved parameters are symmetric. The CBH is not simply an error sink, but is a parameter looked for, taking into account light transmission through the cloud layer. However, we would maintain the main focus of the paper on CTH, and postpone the validation of the CBH product to a later stage.

Addressing once again the issue of low level clouds, given the coarse spatial resolution of GOME, one wouldn't expect an average flat underlying surface over land. Therefore a situation of CTH 1 km $\pm$ 50% can easily lead to unrealistic retrievals. An operational lowest threshold for CTH has been set at 0.8 km. Improving the spatial resolution of the instrument will be advantageous.

> Likewise, what about cloud optical thickness? Have the optical thicknesses been compared with any other standard data sets such as MODIS? More discussion on this should be given.

(A.4) The optical thickness retrieval method of SACURA has been applied to MODIS Terra measurements and intercompared with two other different algorithms (ATSK3 from JAXA and MOD06 from NASA, both based on a LUT approach). Details are given in Nauss et al. [2005]. SACURA COT retrievals exhibit a slightly higher mean than MOD06 collection and ATSK3 (18.5 versus 15.9 and 16.9 respectively) and deviate $\pm$ 18% on average from MOD06, with a stability index r^2 of 0.99. Since the intercomparison has been performed on the same measurement set, the arose discrepancies among the algorithms rule out co-registration and scenario issues and can be tracked down to the different theoretical and algorithmic approaches.

> More discussion on the ground-based data is needed. CTH itself is not a well-defined quantity as the authors acknowledge; IR instruments are sensitive to the radiative height which is not the same as what a lidar measures (see also Menzel et al., 2008). Lidar will not penetrate through a thick cloud. A millimeter-wave radar has more sensitivity in optically thick clouds than a lidar and sensitivity depends upon wavelength. At what wavelengths are these radars operating? In the comparison with ground-based data, which radar is used (Fig. 4)? The number of overpasses selected is quite small. Are these the only overpasses available? How was the selection made? A larger sample size would provide more confidence in the retrievals.

4

(A.5) The original dataset has been presented in Sayer et al. [2011]. The maximum distance for an ATSR retrieval from the ground-based station is 2 km. The coarse GOME pixel additionally shrinks the number of functional overpasses. The ground-based measurements satisfy the quality flag 0-2 (as reported by the ARM document for ARSCL value-added products), this being CTH measured by the MMCR radar. However, with the new set of filters, we have redone the comparison for both deep (N=33 → N=51) and shallow (N=11 → N=15) clouds (as described in the paper), resulting in an increased number of matches. We have also added ROCINN for the corresponding pixels.

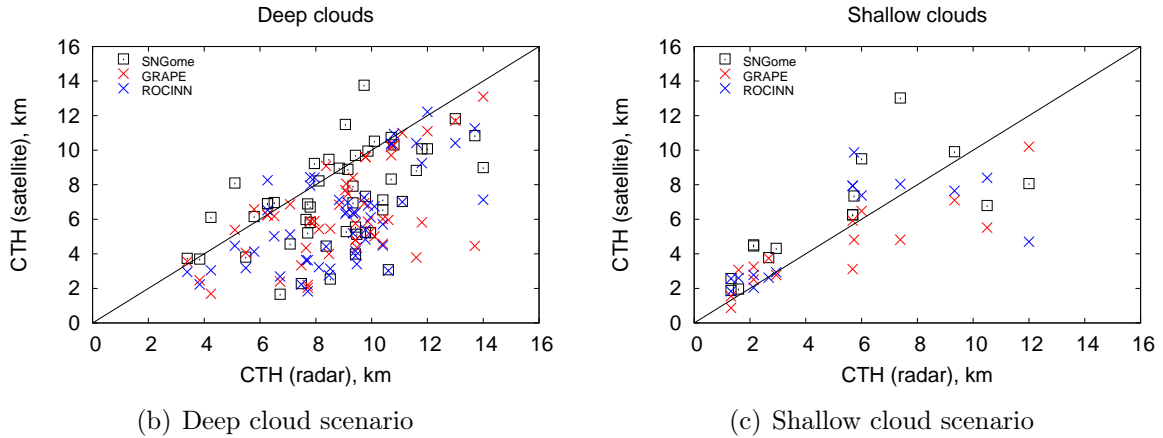(b) Deep cloud scenario   (c) Shallow cloud scenario

Figure 2: Scatterplots for "deep" and "shallow" clouds (as defined in Sayer et al. [2011]) between ground-based and satellite retrievals.

In the shallow cloud case plotted in Fig. 3, the outliers originate from the TWP-Nauru facility. From the climatological viewpoint, this site exhibits frequently westward downwind cloud trails [Henderson et al., 2006], which are, in turn, linked to aerosol production. It is therefore likely that, on the GOME pixel scale, the assumption of a single-layer cloud is not appropriate.

As an example, the radar reflectivity profile for the day 05/07/2001 has been plotted. Given a mean wind speed of 5 m/s and westward direction, the scene sensed by GOME is highly heterogeneous (see Figure 4). We see three distinct layers. At the overpass, the radar CTH was 7.4 km, this being the intermediate layer. SNGome CTH was 13.02 km, for a COT of 10.26. GRAPE placed the cloud at 4.82 km, with COT 2.2, which is the layer of radiative CTH. Clearly the uppermost layer was retrieved by SNGome, handling the space between layers as if it were a single cloud slab.

In the next table, the statistics have been recalculated.

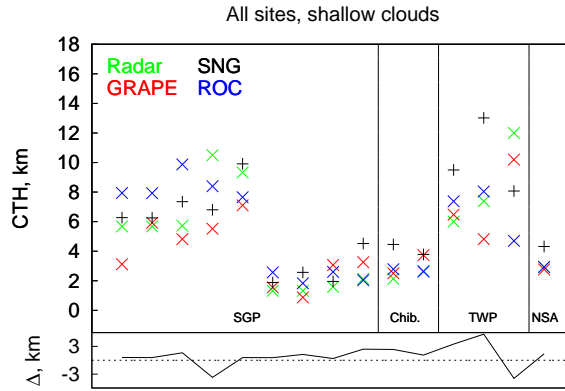|	deep (r)	shallow (r)	Bias (deep / shallow)
Radar	8.51	4.78	-
SNGome	6.89 (0.57)	5.67 (0.75)	-1.62 / 0.89
GRAPE	6.03 (0.52)	4.11 (0.88)	-2.48 / -0.67
ROCINN	5.72 (0.62)	4.96 (0.69)	-2.79 / -0.18

Figure 3: Upper panel: comparison of retrieved CTH as function of ground-based facility for the shallow cloud case of Fig. 2(c). Lower panel: CTH bias between SNGome and Radar.
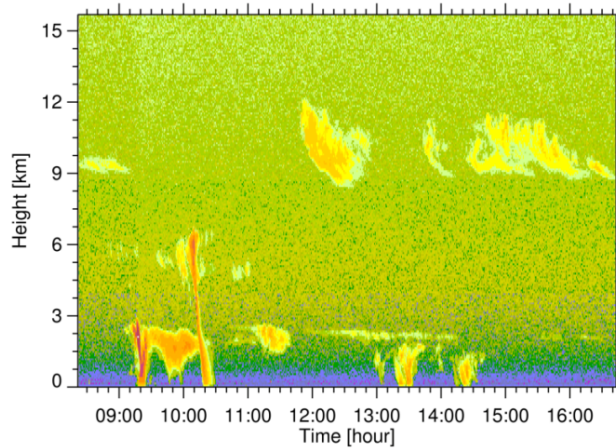

Figure 4: MMCR reflectivity profile of 05/07/2001 at TWP Nauru.

The overall poor agreement between ground-based measurements and GRAPE is disconcerting. The GRAPE algorithm should work well for the deep cloud scenario. We have looked at many MODIS cloud-top pressure (CTP) retrievals (based on the CO_2 slicing approach) over deep clouds as compared with nearly coincident CloudSat data and we see no such errors. Perhaps a comparison should also be made with Terra MODIS which orbits with local time near 10:30. Exact coincidence is not necessary for separate comparisons with ground-based data.

(A.6) Firstly we stress that the distinction between "deep" and "shallow" clouds does not refer to deep convective systems, but it just emphasizes the vertical heterogeneous extent of the sensed scene, as it can be seen in Fig. 9 (shallow clouds) as compared with Fig.10 (deep clouds), p. 3924–3925 of Sayer et al. [2011]. This terminology has been adopted here for consistency among the two papers.

Secondly, we don't think the comparison with GRAPE is discouraging. It is certainly true

6

that the vertical photon penetration depth can be more of an issue with a "deep" cloud. However, in Sayer et al. [2011] a first-order correction for this effect was introduced and resulted in a better (smaller bias) comparison.

Overall we expect to be able to better discriminate different cloud types with more data available.

> There are differences between SACURA CTH distributions (Fig. 8) and those of lidar, lidar/radar, and thermal IR shown in Stubenrauch et al. (2010). The lidar/radar measurements were screened for subvisible cirrus. Stubenrauch et al.'s plots go to zero-1 km and show a large fraction of clouds at low altitudes for a wide range of latitudes. In contrast, SACURA retrievals do not show high fractions of low clouds (below 5 km) in the northern hemisphere either in winter or summer and do not show many high tropical clouds in the boreal winter. More discussion is needed here.

(A.7) We show here new plots with the updated screening procedure. The bin width has been increased by a factor of two and the scale range now is [0–15]%. The general structure of clouds is preserved. The features resemble AIRS-LMD distributions, especially in the tropics for high clouds, where SNGome places the maximum at $\approx$ 13 km in boreal winter, and in the southern mid latitudes for low clouds. This behavior is expected, because in the case of a thick layer underneath a thin one, both sensors detect the former.
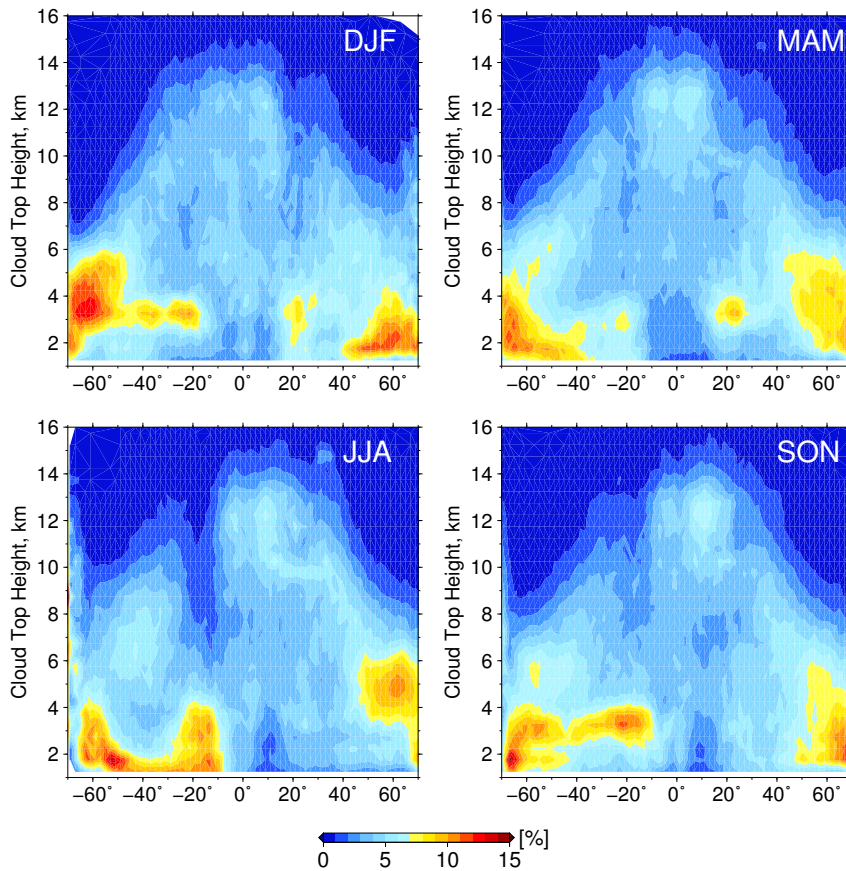

Figure 5: Seasonal CTH occurrences for year 2001

7

The CTH PDFs over land in Fig. 11 show a unimodal distribution. Please see references in Joiner et al. (2011) that suggest/show distributions that are bimodal over land. Again, more discussion is needed.

(A.8) For the same year 2001, the disentangled plots have been reanalysed. The results show more features. For instance, given the global land/ocean ratio, comparing bottom-right plot (southern hemisphere, over water), the PDFs suggest 3 modes at $\approx$ 1.8 km, 3.7 km and 5.5 km and resembles the upper plot of Fig.13, p. 6225 in Joiner et al. [2011] for the CLOUDSAT profiles. Similarly, over land, SNGome shows a bimodal PDF. A filter for CF < 0.3 has been applied, in order to screen occasional dust events and to be consistent with the plots in Joiner et al. [2011]. Below, the global PDF for year 2001 is reported as well.
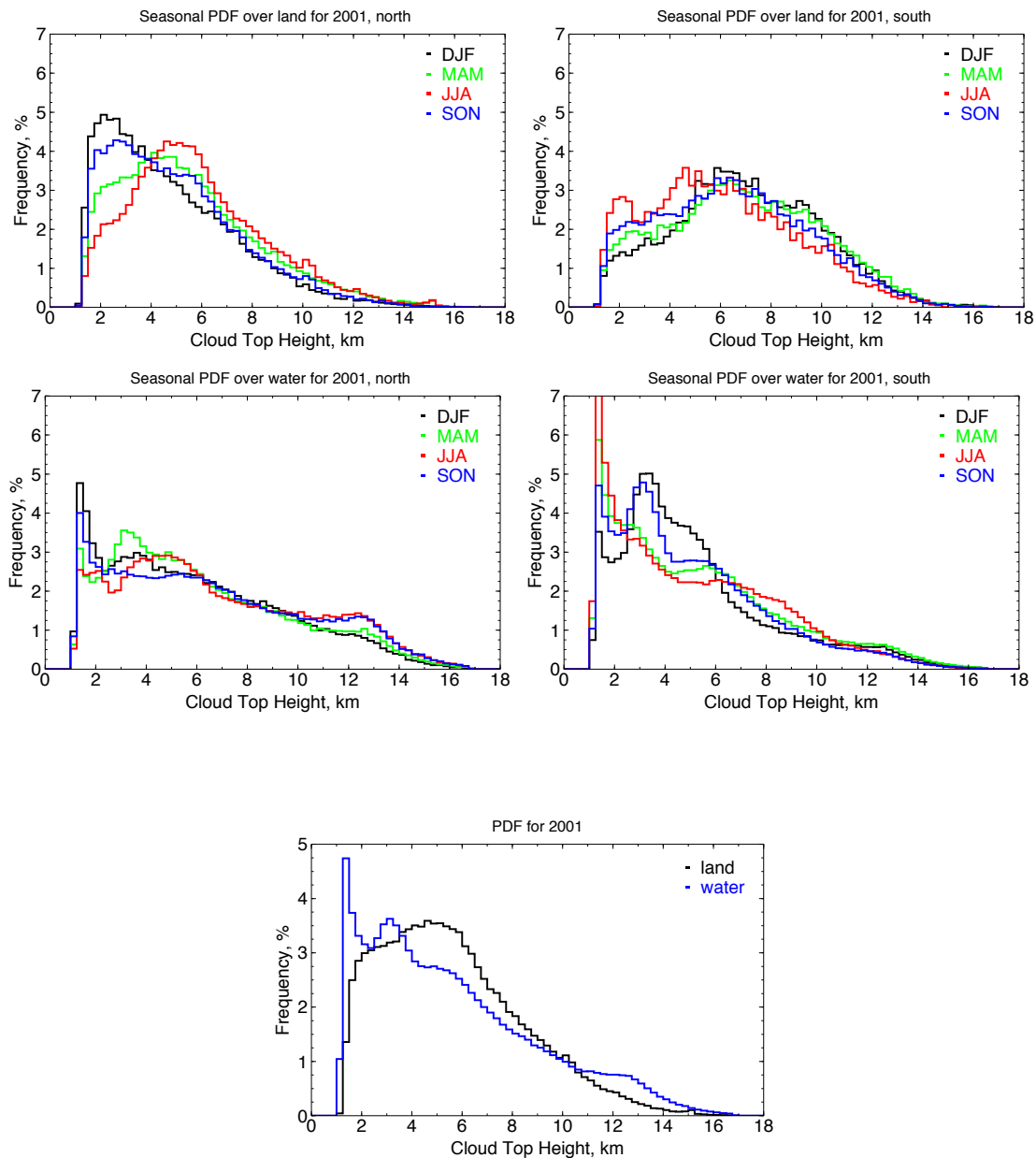

Figure 6: Seasonal and annual cloud top height distributions for year 2001 over land and ocean, disentangled into northern and southern hemisphere, from the SNGome dataset.

8

Specific comments: In Fig. 2, why do the glory and rainbow effects occur only for a limited range of CTH (i.e., not below 6 km and for some angles not above 10 km)? In Fig. 3, why do the SZAs where the errors peak not match the glory and rainbow angles where the CTH errors peak in Fig. 2? Is there an explanation for the large errors at high SZA?

(A.9) Given the geometry of the experiment of Fig. 3, a SZA $= 38°$ corresponds to a scattering angle of $142°$. Referring to Fig.3.7, p.152 in Kokhanovsky [2006] (reported below for convenience), we see that the feature is explained by the presence of rainbow. For additional explanation of Fig.3, please see point A.18 in the response to referee #2.
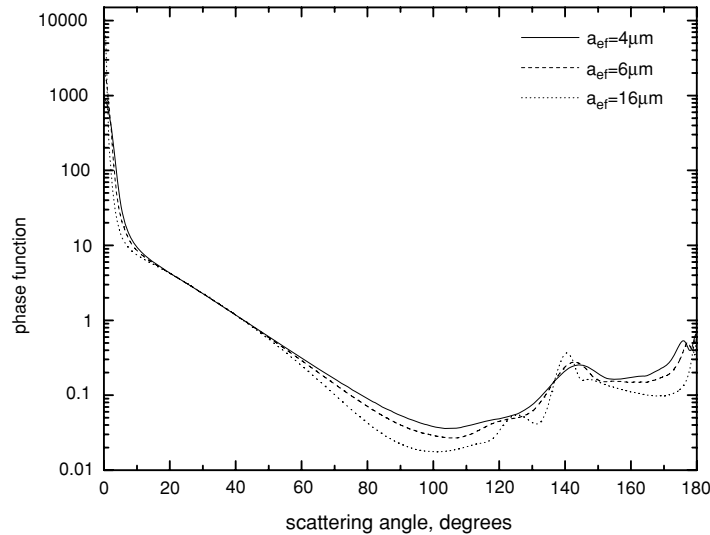

Figure 7: Dependence of the cloud phase function on the effective size of water droplets at λ=0.55 µm. The Gamma PSD with radius 6 µm is used in the calculations.

On p. 5000, it says that only a quality flag of 5 (best convergence) is used. What is the meaning of other values of the quality flag and how does choice of this quality flag value affect the results?

(A.10) The algorithm flags each retrieval in ascending order, depending on the quality of the simultaneous fits of CTH and CBH, given the value of COT calculated in the continuum outside the band. In the first draft only retrievals flagged 5 were used (i.e., the gray curve in Fig. 1).

Value	Description
0	No retrieval
1	Only CBH
2	Only CTH
3	Geometrical thickness limit
4	No convergence
5	CTH and CBH convergence

In this reanalysis, retrievals flagged 0,1 and 4 are clearly discarded. In view of the two-layer study presented above, we have used retrievals flagged 2,3 (given a CTH < 5 km) and 5. The use of flag 2 is justified by the fact that the main focus of this paper is CTH and, only in this case, CBH works as an error sink.

> Figs 4 and 5: Would perhaps be better to show differences as a function of cloud fraction and/or cloud optical thickness.

(A.11) We address this with the next answer. Please see also response to referee #2, point A.35.

> Fig. 6: There are significant differences between SNG and ROC, sometimes positive, sometimes negative, sometimes quite large. The discussion does not explain all the differences. More discussion/analysis would be helpful. The bottom part of Fig. 6 is hard to see. Suggest breaking it out and showing with larger vertical spacing so that differences (which appear substantial) can be better seen.

(A.12) It was believed that, due to the great GOME pixel size, an underestimation of cloud fraction would have impacted the top altitude retrieval, via the Independent Pixel Approximation. So we would have expected a correlation between the quantities plotted in the panels. It is not the case. Assuming the ATSR CF as the true one (due to the better spatial resolution) we see no effective correlation. Please see also point A.35 in the response to referee #2.

> Technical corrections: p. 5006, L 13: typo, negative should be negatively

(A.13) Corrected.

Bibliography

B. G. Henderson, P. Chylek, W. M. Porch, and M. K. Dubey. Satellite remote sensing of aerosols generated by the Island of Nauru. *Journal of Geophysical Research*, 111(D22), November 2006.

J. Joiner, A. P. Vasilkov, P. Gupta, P. K. Bhartia, P. Veefkind, M. Sneep, J. de Haan, I. Polonsky, and R. Spurr. Fast simulators for satellite cloud optical centroid pressure retrievals, 1. evaluation of omi cloud retrievals. *Atmospheric Measurement Techniques Discussions*, 4(5):6185–6228, 2011. doi: 10.5194/amtd-4-6185-2011. URL `http://www.atmos-meas-tech-discuss.net/4/6185/2011/`.

A.A. Kokhanovsky. *Cloud Optics*. Springer, Dordrecht, 2006.

A.A. Kokhanovsky and V.V. Rozanov. The physical parameterization of the top-of-atmosphere reflection function for a cloudy atmosphere-underlying surface system: the oxygen a-band case study. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 85:35–55, 2004.

L. Lelli, A. Kokhanovsky, V. Rozanov, and J.P. Burrows. Radiative transfer in the oxygen A-band and its application to cloud remote sensing. *Atti della Accademia Peloritana dei Pericolanti - Classe di Scienze Fisiche, Matematiche e Naturali*, 89(S1):C1V89S1P056–1–C1V89S1P056–4, 2011.

T. Nauss, A. Kokhanovsky, T. Nakajima, C. Reudenbach, and J. Bendix. The intercomparison of selected cloud retrieval algorithms. *Atmospheric Research*, 78(1-2):46–78, November 2005.

K. Sassen, Z. Wang, and D. Liu. Global distribution of cirrus clouds from CloudSat/Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) measurements. *Journal of Geophysical Research*, 113, October 2008.

A. M. Sayer, C. A. Poulsen, C. Arnold, E. Campmany, S. Dean, G. B. L. Ewen, R. G. Grainger, B. N. Lawrence, R. Siddans, G. E. Thomas, and P. D. Watts. Global retrieval of ATSR cloud parameters and evaluation (GRAPE): dataset assessment. *Atmospheric Chemistry and Physics*, 11(8):3913–3936, 2011.