Reply to comments by reviewer #4:

(our replies are boldfaced)

Reviewer 4: amtd-8-C2898-2015

Introduction: The authors present a fairly nice listing of existing cloud measurements. I miss an assessment of cloud height measurement quality/accuracies from the different instruments, though, as well as a review of the requirements on cloud height accuracy for current scientific problems.

Q4.1: Are, e.g., CTHs with 1.5km or worse accuracies (as yours are by definition) valueable?

- **A good question that I might not answer. A single measurement with such a coarse accuracy is nowadays not valuable. The high precision instruments like CALIPSO can do much better. Global measurements for nearly 10 years on the other hand can be very useful for verification purposes, the detection of sources like volcano eruptions, or helping to better constrain trace gas retrievals towards the lower tangent heights.**

Q4.2: Also, a more concise overview of the strenghts and weaknesses of the different instruments would be nice, leading to what SCIAMACHY limb measurements offer beyond what other sensors provide.

- **This seems to be beyond the scope of this paper. As already mentioned in the previous question our aim is to highlight the strengths and weaknesses of the retrieval by giving examples like differences wrt nadir data or detecting aerosol layers after the major eruptions.**

Q4.3: Data: State clearly, which data is used for what: SCIA limb for your original retrievals; SCIA nadir and MIPAS for comparison/validation. Which versions of the comparison data have been used? Provide references.

- **Done.**

Method: The approach uses a two-wavelength (color index) ratio. I find the argumentation how the wavelengths bands are chosen insufficient as well as the explanation of cloud signatures in the CIR. Questions like these remain unanswered:
Q4.4: why these 2 bands exactly?

- **Already addressed in Savigny (2005).**

Q4.5: or are any two bands ok?

- **Any band that is not influenced by molecular absorption and instrumental problems like dead/bad pixels can be used. We are using other bands as well (see Fig.10). But results are not much different except for the NIR band that takes ice absorption into account. This will help differentiating between aerosol layers and clouds in the upper troposphere and is work in progress.**

Q4.6: why 2 at all?

- **For a ratio technique, 2 wavelengths/bands are enough.**

Q4.7: Why is at tangent height radiance ratio at one wavelength not sufficient?

- **It might be sufficient using a different retrieval technique**
- **Normand, E. N., Wiensz, J. T., Bourassa, A. E. and Degenstein, D. A.: Cloud discrimination in probability density functions of limb-scattered sunlight measurements, Atmos. Meas. Tech., 6(12), 3359–3368, doi:10.5194/amt-6-3359-2013, 2013.**

Q4.8: what causes the different tangent radiance ratio behaviour in the two different bands?

- **Wavelength dependent solar radiation is scattered.**

Q4.9: What CIR pattern have non-cloud cases, what patterns are introduced by clouds (an example of non-cloud CIR profile would help the understanding)?

- **See Savigny (2005), Fig.3a : There is no pronounced pattern, CIR is near 1. See also the added CI/CIR cloud free case in Fig. 5 of the manuscript.**

Q4.10: What causes the peak (is it a peak at all, i.e. is CIR decreasing again after a cloud? under what conditions?)?

- **I think this all can be seen in Fig. 5. Due to ratioing tangent heights we get a positive peak when CI sharply drops with height.**
- **(1) More scattering occurs due to extra particles in the line of sight of the instrument and thus more radiance enters the instrument. (2) Particles scatter differently wrt wavelength changes than molecules.**

Q4.11: How has the threshold been chosen (no information at all on this is included in this paper, though i assume, that is is a fairly crucial choice)?

- **The explanation was added to the manuscript. Enhancing the threshold (1.5) does not have a significant effect on average values (see supplement Fig.)**

Q4.12: Simulations: I find the sensitivity analysis presented too sparse to judge the quality of the method. Particularly, later in the paper the authors draw conclusions from sensitivity checks that have not been presented, i.e. which are intransparent and not reproduceable by the reader.

- **We added more sensitivity tests to better cover the natural variability of clouds.**

Q4.13: I like to see more details on the sensitivity tests including systematic teanlysis of cloud height dependence of the CIR (which was obviously done and argued with later, but the reader gets presented only a selective summary, which hinders to see and understand the bigger picture), tangent height dependence, and effect of higher clouds in low tangent height measurements (clouds that might be out of FOV in the high(er) tanh-h observation; that such cases can occur is mentioned in sec3 and discussed in sec4.1, but their appearance in a CIR profile and the sensitivity of the measurements to these is never analysed). Preferably also sensitivity to cloud “concentration” (extinction, geometric thickness – cloud with same _ likely gives different signal depending on its geometric thickness. but how different is it?

- **A figure on CIR vs. CTH was added to the supplements.**
- **The effect of concentration (optical thickness) was done by choosing in the examples of Fig. 6 only very low COTs, so that any increase in COTs will only enhance the CIR.**
- **The influence of geometric thickness was tested and the CIR was insensitive to it. We also have chosen only clouds of 1km thickness for the model studies to show that thin cloud layers are not a problem for the method.**

Q4.14: how munch cloud needs to be within the FOV to get a detection?),

- **Very small amounts of clouds need to be in the LOS, in the best case COT=0.003 is enough.**

Q4.15: cloud microphysics (specifically, size of ice particles).

- **It was tested again and the effect of changing particles size was negligible for CIR (1%)**

Q4.16 Then, I expect stringent conclusions: where are the detectability limits in dependence of cloud height and cloud thickness?

- **The higher the cloud, the better the detectability. Limits are given in the text.**

Q4.17 Also: how sensitive are these conclusions to the threshold choice?

- **The threshold is not a problem for higher clouds as the CIR is much higher. See fig.6, even for thin clouds and moderate aerosols the CIR is larger than 1.7.**
- **The main problem is found at lower tangent heights where AOT needs to be low. A decrease in threshold would lead to more cloud detections at wrong, higher heights because of the influence of aerosols.**

Q4.18 Maybe even: how many clouds are expected to be missed at the different altitudes?

- **This is hard to answer but comparisons with MIPAS showed that at 8km the differences start to rise. This is due to detection sensitivity differences from two different wavelength regions used for the retrievals.**
- **At 3 km is would expect to miss the most of the clouds, but we still detect a considerable amount of low clouds in favorable areas (e.g. with very low AOT)**

Q4.19 Results: For a measurement that is limited to provide one cloud detection per profile, height resolved cloud fractions seem quite a complex (and somewhat confusing) parameter to look at. On the one hand no cloud hits, i.e. no contribution to CF, will happen below a detected high-altitude cloud, which leads to a shadowing of lower clouds (lower low-cloud fractions in regions with lots of high clouds).

- **Correct.**

Q4.20 On the other hand, high clouds that have been outside the FOV of the high tangent altitudes but are in the FOV of a low tangent altitude are assigned to the low tangent altitude, both in altitude as in geographic position (displacement) and contribute to cloud fraction in the low layer. I find it hard to wrap my head around that and wonder whether the authors were sucessful to do so

- **Due to averaging over a year in our example these effect are cancelled out.**

Q4.21 In the time trend discussion, different versions of the 2-band CIR are aplied. However, as characteristics of different band selections are never presented or discussed, the reader hardly can follow the argumentation and conclusions, which makes the whole discussion pointless.

- **Now only one version used.**

Q4.22 Validation: Most of all, I wonder, why no comparison or validation against CALIOP data is presented. CALIOP by far provides the best dataset to compare to when promising to be able to provide thin cloud CTH. Comparison to SCIA nadir and MIPAS CTH is nice, but rather underlines the issues that each of these datasets has (nadir only sensitive to thick clouds, MIPAS – as SCIA limb – just to uppermost cloud layer CTH and with cloud geolocation issues).

- **Yes, a comparison to CALIPSO would be necessary for an extended validation.**
- **This paper was written to show the technique and give a few examples for a cloudy atmosphere.**

Q4.23: The modification of the SCIA nadir CTHs with their standard deviations seems fairly random. I miss a rationale for taking the standard deviation (why not just taking a constant offset? Or a height dependent offset?). Also, original nadir data are never presented (neither mean nor standard deviation), so that the reader can not judge at all the effect of your – random – modifications. Nevertheless, the never presented data is discussed ("Without the added standard deviation, nadir CTHs were systematically and globally lower. . . . ".

- **This will be added to the supplements.**

Q4.24: The authors claim a good agreement of SCIA modified nadir and limb CTH without providing sufficient evidence. From comparing Figs7 and 12, I do not necessarily conclude a good agreement. Basic patterns are quite different (e.g. very high clouds off the western coast of middle America in nadir CTH; fairly high nadir CTH over all of South America, Southern Africa, and Australia). Any quantitative comparison is miss-ing, although opportunities for these two measurements are better than for any other (there should be plenty of collocations!). I like to see e.g. a scatter plot like presented for MIPAS CTH and some difference statistics (mean diffs, std of diffs e.lg. for different regions, cloud regimes, …).

- **The paper was made to publish our retrieval model as it is used in operational Level 2 Version 5 data. Thus we do not intend to make a full validation but to highlight the main results so far.**
- **For an exclusive validation paper the point mentioned should be addressed.**

Q4.25: In Sec4.1. horizontal sampling issues have been disacussed in quite some depth. It appears in the discussion of the SCIA limb CTH themselves (Sec5), but is not

considered anymore in the comparison to other CTH data, although I expect it to play a significant role.

- **Comment added.**

Q4.26: Volcano event: This whole section seems a little disconnected from the rest, both in terms of target and style. I wonder, whether this would rather suit a different or separate paper.

- **It is added because we had the 3 major volcano events in our data and we analysed how good we were catching these events compared to literature data. Compared to MIPAS we have found some discrepancies due to detection sensitivity.**

Q4.27: Throughout the paper, colloquial language (or lab slang) and imprecise formulations is ocassionally used (examples see below). Please reformulate properly. This could improve perceptibility of the paper a lot. References are used in odd manner at times (e.g. Reuter&Pfeifer; von Savigny et al., 2005; Hommel et al.). Use appropriately, replace with proper ones, or remove.

- **Partly removed.**

Generally, I have the feeling that too many discussion items and conclusions are merely stated, but not backed up sufficiently with explicitly and clearly presented evidence.

2 Specific comments

Q4.28: p8296, l20: what does “global cloud field distribution” mean?

- **Rephrased: global distribution of cloud top heights**

Q4.29: p8297, l11: “Clouds occur in complex shapes” – shape as in contour/outline? if not, can you find less ambiguous wording?

- **Rephrased**

Q4.30: p8298, l1f: “The balance between these effects” – which effects? albedo enhancement and cloud emission are obviously not in balance.

- **Rephrased: compensation, as in**

p8298, l3: “The light scattering ability of cloud droplets is only weakly wavelength dependent” – This is at best imprecise. Wavelength dependence in Mie regime is weaker than in Rayleigh regime, and in VIS/NIR cloud droplets are in Mie regime, while molecules are in Rayleigh. The picture changes, however, when going to other spectral regions.

- **True, added wavelength range**

p8298, l7: “This also enhances light absorption by trace gases [. . .] above the clouds.” – Above? Why is that?

- **Corrected, above removed.**

p8298, l15f: “Due to the low temperatures [. . .], cirrus clouds are made of small ice crystals.” – It’s not the low temperature, but the low absolute humidity.

- **Due to the temperature at their altitude of formation, cirrus clouds are composed exclusively of ice crystals (Cziczo, 2013).**

p8299, l18f: “Nadir viewing, passive instruments [. . .] were mainly designed to derive trace gas columns” – I do not agree. Some yes. But many more were designed for weather observations (including clouds).

- **That’s true. Corrected.**

p8299, l25ff: “Based on the [. . .] (ISCCP) D2 dataset, a global COD _N = 3.9 ± 0.3 was found [. . .] showing that a lot of clouds types cannot be retrieved.” – What type of instruments is the dataset based on? How does global mean COD show that many types (types? others than cirrus?) can not be retrieved? Or even that many cloud occurences can not be retrieved? I’d like to see a more stringent argumentation.

- **The infrared and visible radiances obtained from imaging radiometers carried on the international constellation of weather satellites were used, see Rossow (1999)**

p8299, l27: “with a 19.6% global coverage” – reference missing (or is it from R&S, 1999, too?).

- **Corrected with newer data from Sassen (2009).**

p8300, l3: “An average global cloud amount of about 68% was reported” – Rather cloud coverage? Amounts can hardly be measured as percentage (what is the reference?).

- **Maybe. Phrase taken from Stubenrauch (2013), reference added.**

p8300, l25ff: Revise the paragraph. Make clear from the beginning that now, it is SCIAMACHY limb measurements and cloud retrievals that you talk about. “SCODA was developed using the colour ratio method” – there is not only one, well-established (aka “the”) colour ratio method.

- **Yes, corrected**

p8302, l3ff: Mention overpass time. What is the horizontal spacing of two limb sequences?

- **7° which is roughly 780km. information added.**

Q4.40: p8303, l19ff: The whole paragraph probably fits better in Sec3. In any case, more stringent explanations are required for the reader to be able to follow you. E.g. “Higher intensities can be seen in channel 6 when cloud scattering occurs” – higher than what? are they really restricted limited to ch6? “A steeper gradient between radiances from both channels” – i.e., I15.4km
ch4 /I15.4km
ch6 > I12.1km
ch4 /I12.1km
ch6 ? “This means that clouds are in the field of view at tangent height 12.1 km and scatter more radiance into the line of sight.” – This does not explain the (changing) “color” gradient. So, where does the color gradient come from? And why does a changing or high gradient imply a cloud? What are the physical processes causing that? Would be helpful to see a Fig.2 equivalent for a cloud-free case.

- **Revised the paragraph, changed figure.**

p8303, l26ff: What is the point of this paragraph? Discussing the choice of spectral retrieval windows? It is insufficient for this. And does not present any conclusions/ choices.

- **Connected to the part where the channels are described.**

p8303, l26: “Spectral windows with molecular absorption bands” – Reformulate, Common terminology is that (spectral) windows are the regions with low gas absorption.

- **Done.**

p8304, l12f: Does tangent height knowledge refer to geometric heights, or are refraction effects included?

- **Geometric. Added to text.**

p8304, l24ff: Why/how is limb/nadir matching (and what is that?) related to platform height?

- **Added reference (gottwald, 2011).**

p8306, l15: “The standard operational CI approach” – what is the standard approach”? provide a reference.

- **Done**

Fig.3: Wavelength bands indicated by dotted, not dashed lines (as claimed by caption).

- **Corrected**

p8307, l1: “Water and ice clouds show similar reflection characteristics in this wavelength range but start to differ at wavelengths larger than 1400 nm” – Which characteristics specifically are meant? Your analysis later on (p8312) seems to contradict this statement.

- **Rephrased.**

p8307, l18: “where “h” denotes the high and “l” the low band.” – high and low what? wavelenghts?

- **Yes, corrected**

p8308, l6: “Figure 5a shows an example of the cloud index ratio profile” – Fig5 seems inconsistent with Eq. (1) and Fig2. The latter indicate CI<1 (Ih<Il) and CI closer to 1 for lower tangent altitudes, while Fig5 shows only CI> 1 and CI being large at low tangent altitudes.

- **The normalization by the highest tangent height for CI was used for visualization purposes. Corrected in text**

Q4.50: p8308, l15: “but we only store the highest peak” – ’highest’ as in highest tangent altitude or as in largest CIR?

- **Corrected: highest altitude**

p8308, l27: “the retrieved cloud fractions” – first occurence of ’cloud fraction’. not straight forward, what CF means for limb measurements, hence requires explanation. here. Besides, I find the term cloud occurence frequency more suitable (cloud fraction, in my opinion, is rather the coverage within a sensor FOV or a modeling grid cell, i.e. describes sub-resolution issues).

- **Cloud fraction was changed to cloud detection rate and as a height resolved quantity to cloud occurrence frequency as also used in MIPAS papers on CTH**

p8309, l1ff: What area exactly do you calculate here? The area within the instrument FOV intersecting a _zth layer (that’s what I would expect as the correct one), or the area of a _zth thick plane layer intersecting a _zth thick shell (as Fig.2 suggests)? and to be more correct, shouldn’t it be the intersection volume? would that make a significant difference?

- **Rewritten the paragraph**

p8309, l19: “the air mass above a cloud has less air” – less air than what? in the cloud-free there’s as much (or little) air in this region as in the cloudy case. revise formulation.

- **Revised**

Fig.5: “(see Figure above)” – Refer to specific, numbered figure.
missing: I miss a clear statement that you method restricts CTH retrieval accuracy to tangent altitude spacing, i.e., it is kind of binned data. This should also be considered in the validation discussion.

- **Information added.**

p8309, l25f: “Simulating radiances in cloudy atmospheres is difficult [. . .] because radiative transfer models [. . .] treat clouds only as two-dimensional layers” – First, cloudy simulations are complex even if one fully considers 3D properties, purely due to the complex nature and high variability of clouds themselves (as you mentioned in the Introduction). Second, please avoid making it sound like all models have this limitation (e.g., DART and MYSTIC provide 3D capabilities for scattered solar radiation). Lastly, what do you refer to by ‘two-dimensional layers’? Do you actually take into account layer inhomogeneity in the observation plane?

- **Cloud were modelled with different thicknesses, for the analysis presented we only used a 1 km thick cloud to have a reduced cloud appearance in the FOV. The layer was taken to be homogenous.**

p8310, l12: “A model accuracy of 5% was found” – compared to what? what constitutes the truth/reference? Is the accuracy estimate valid at all with the forward

model reference from 2014, the accuracy reference from 2003?

- **This was corrected as 5% was not state of the art anymore. Now only 2014 reference used, which is much more appropriate.**

p8310, l12: “Cloud top heights can best be retrieved by using the same instrument and measurement mode” – generally? why would that be? or just if used for cloud-flagging in trace gas retrievals (if so, how is it relevant here?)? best for what? what at all is the point to discuss retrievals here again, in the forward modeling section?

- **Yes, rephrased. Moved to retrieval section.**

p8310, l22: “The SAA is the angle between the line of sight and the sun beam” – no, that would be the single scattering angle.

- **Sentence deleted, not necessary.**

p8310, l28f: “The Rayleigh scattering phase function” – aren’t you discussing clouds? which are in Mie regime?

- **Sentence deleted, not necessary.**

Q4.60: p8311, l7ff: “A spherical atmosphere [. . .]. The scalar discrete ordinate technique was used” – _____Does it apply a true spherical or a pseudo-spherical solution for the scattering calculations?

- **True spherical, reference added**

p8311, l13ff: Is the spectral setup consistent with SCIAMACHY’s instrument characteristics and the spectral bands used in the cloud height retrievals? what intermediate angle steps are calculated?

- **Yes.**

p8311, l17: “To find the limitations” – Of what? Cloud retrievals? Trace gas retrievals? . . . ?

- **Cloud retrievals, corrected**

p8311, l20: Do you use monodisperse clouds? Monodispersions are known to exhibit oscillations in their optical properties (scattering efficiency as well as phase function), which are significantly dampened in realistic bulks. Are you sure this affects your sensitivity estimates in a negligible manner?

- **See Rozanov 2014, e.g Fig. 9. For water polydisperse. References added.**

p8311, l23: What is the source of the ice crystal optical property data (assuming the droplet properties are from Mie calculations)?

- **See Rozanov 2014, reference added**

p8312, l6: What is the rationale to use Henyey-Greenstein for aerosols? Based on what have the assymetry parameters been chosen?

- **See Rozanov 2014, reference added.**

p8312, l7ff: “In Fig. 6 [. . .]” – Simulation setup description is redundant. However, for which tangent altitudes have the simulations been made? Do these simulations include the aerosols?

- **Figure corrected**

p8312, l12f: “Cloud tops were in general detectable [. . .], because the modelled Theta exceeds the threshold of 1.4.” – This seems to be the wrong argumentation direction. Where at all does the threshold come from (explanation missing in Sec3)?

- **corrected**

p8312, l19: “The phase function of an ice particle is more complex and has a stronger asymmetry. Thus [. . .]” – This contradicts you statement from p8307, l1.

- **Argumentation changed.**

p8313, l26ff: What is the effect of having the surface in the FOV on CIR? over the

range of possible surface reflection properties (is sunglint an issue?). can the cloud detection threshold be reached in non-cloud cases due to surface effects? considering this, is cloud detection reliable for views with surface in FOV?

- **Sunglint is no issue due to the SAA/SZA angle combinations**
- **Towards the surfaces due to your mentioned effects, the method has its limits.**
- **The ocean in this wavelength range is rather dark.**

p8314, l5ff: “But the cloud retrieval can be obstructed [. . .], when the optical thickness of cloud and aerosols has the same order of magnitude.” – Do you say that a thin aerosol layer (say _=0.01) would not obstruct a thick low water cloud (say _ =5), since they do not exhibit same order of magnitude _? Reformulate otherwise.

- **Yes, as shown in Fig. 6a, even with low AOT = 0.07, we detect the thick cloud at the right height**.

Q4.70: p8314, l8f: I suggest to make it clear already in the title, that along-track, or rather along-LOS, sampling (resolution?) is discussed here (in contrast to the across-track sampling that SCIAMACHY is doing).

- **Good suggestion, done.**

p8314, l10ff: That real sample path is longer than the 1-shell path 2x(i,i) is NOT an effect of the 2x(i,i)-shell not being resolvable. Reformulate.

- **TODO**

p8314, l15: I have problem to follow. Where is tangent height going into this? A terminological distinction between true CTH (wherever along LOS) and retrieved CTH (at tangent height) could be helpful.

- **corrected**

p8314, l24: “For most of the cases [. . .]” – Do you mean that for no ‘tomographic LOS intersections’ exist within the troposphere?

- **Subsection rewritten**

p8314, l26ff: “But if the LOS of a cloud free pixel [. . .]” – Why would the resolution be depending on (non-)cloudiness?

- **Subsection rewritten**

p8315, l9: “to derive a trace gas profile at the tangent height.” – ‘profile’ commonly refers to vertical dependy of a parameter (e.g. VMR(z)). How do you relate it to a tangent height, which is just 1 point in z? Or do you mean the profile at the (horizontal) geographic position of the tangent point?

- **corrected**

p8315, l11ff: “the retrieved CTH will be higher in this case. These high clouds outside the along-track FOV or lowest shell can lead to a CTH overestimation” – CTH will be higher compared to what? why would a cloud outside the FOV at all lead to a wrong estimation of CTH? I do not understand how any of this could lead to a CTH OVER-estimation. The only one, I see, is an underestimation: if there is a cloud high along the LOS, which is not in the FOV for the corresponding high tangent altitude. This high cloud would cause a cloud hit for the low tangent altitude, i.e. a low CTH estimate.

- **Outside the FOV was wrong, rewritten.**

p8315, l13: cloud fraction needs to be introduced before discussing it.
Fig. 6: “Simulated maximum colour index ratios” – Maximum over what? And why, maximum at all? Furthermore, what tangent altitudes are those figures for? what does ’T_A’ in plot titles refer to?

- **COF introduced and figures updated**.

p8315, l14: Surely a matter of taste, but Sec6&7 are results, too. Hence, I’d chose a more specific Sec5 title.

- **Changed to Retrieval Results**

p8315, l16: "We retrieved annual means of cloud top heights and fractions" – 'Annual mean cloud fraction' does not make sense. Formulate properly.

- **I don't understand the problem. But fraction was changed to COF**.

Q4.80: p8315, l16ff: Cloud fraction should be a relative number, not a pixel count. Formulate properly. Furthermore, how is the geographical position assignment done? Is the cloud assigned to the position of the tangent point?

- **Changed**
- **Cloud were assigned to the tangent point**

p8315, l21: Any rationale for the grid size? As you have demonstrated before, the path lenght through a CTH shell is about 400km, i.e. roughly 4_ big. Seems more resonable to me to have the grid big enough such that all of the possible cloud positions within this shell fall into one cell.

- **Good enough for annual means. 400.000 measurement per year**.

p8315, l21f: "All measurements [. . .], that exceeded the threshold" – Here, one limb scan is one measurement? Or does it include all points within a vertical scan that exceed the threshold?

- **Rephrased**

p8316, l1: "A cloud image (Reuter and Pfeifer, 2011)" – what is the Reuter&Pfeifer contribution to this image, what is special about it that it is referecned? particularly, here in text instead of in figure caption (e.g. as 'adapted from')?

- **Reference corrected, Reuter supplied the image.**

p8316, l12f: "which will be analysed later." – future work? or in this paper? if the later, specifically refer to the section, where it is discussed.

- **corrected**

p8316, l16: "This is partly due to dependence of the phase function on the viewing geometry" – what else is it/could it be due to? Phase function (as an inherit property of a particle) does not depend on viewing geometry. Single scattering angle does (and hence, 'efficient scattering coefficient' = scat.coef * pfct(SSA)).

- **Corrected, added another effect: stray light**

p8317, l4f: "This can be attributed to" – Is it absolutely clear that this is the reason, or is it speculation (if so: might be attributed)? Could smearing (or displamenent along LOS) effects play a role, too?

- **Limb smearing can not be ruled out but it means that we need higher clouds in the north of the zone. This would only influence a small part northwards. So that seems negligible.**
- **Content with reference added.**

p8317, l13: "Clouds are limited [. . .]" – what's the point here? i don't understand what you are trying to tell the reader. Besides, are PSC not clouds (as they are limited to troposphere)? Use a better reference for the general statement on PSC altitudes (this fact has been known far before the von Savigny paper).

- **Paragraph deleted**

p8317, l20: "It seems that the method is rather insensitive to low clouds" – You never mention or discuss the "shadowing" effect of high clouds (only highest cloud in a limb scan is detected, clouds further below are neglected. This will produce a low bias of low clouds in regions with frequent high(er) clouds. Might this contribute? Also, might displacement effect contribute? Particlularly, displacement effects could also have a directional dependence (i expect that looking from a high-cloud region into a low-cloud region along track (or LOS) gives a different signature than looking from low- into highcloud region).

- **That is reasonable and can happen. The shadowing effect was discussed in Appendix 1.**

p8317, l24: What sensors did Generoso et al. use?
- **Calipso, mentioned**

Q4.90: p8318, l1: "But CALIPSO [. . .]" – CALIPSO (or CALIOP? as you made that distinction between platform and instrument before) should also be able to detect enhanced aerosol layers. That is, this can not be the reason for not appearing in CALIPSO. Did you check CALIPSO aerosol data? What else can explain the absence of a signature in CALIPSO? Without another explanation, I'd rather suspect it to be an artefact of your method.
- **Using CALIOP/CALIPSO now.**
- **Aerosol data checked and reference made**

p8318, l25f: "aerosol layers are optically very thin and the corresponding CIR is lower than for typical tropospheric clouds." – is this due to low _ or due to the aerosols as such (a stringent presentation of cloud altitude/optical depth dependency of CIR in Sec3 would help greatly)?
- **Corrected**

p8318, l27: "Two other wavelength pairs were tested." – Which? Why and why these in particular? How were the wavelengths selected? What signature do we expect there? Is 1.6/1.5um combi always constant, or just its mean?
- **The two other pairs are also retrieved in our processing. The 1090/1550 ratio is used in the water vapour retrieval. The 1.6/1.5 ratio was originally used to detect ice clouds but the difference to the standard ratio was negligible as most cloud tops are icy. But it can be used for discrimination of clouds and aerosols in the upper troposphere. This is ongoing work at the moment.**

p8318, l28: "During the longer decontamination phases" – Does this imply some instrumental issues affecting the CIR retrievals?
- **Yes, as the temperature of the detectors is raised**.

p8319, l3: "This indicates that the ratio is not affected by aerosols as much." – Which of the ratios? Not as much as what? Is there a (physical/optical/rad.transfer) explanation for that? You get the reader far too little background/explanation to follow any conclusions you draw (1.6/1.5um combi mean is so stable that I suspect it might not show any cloud signature at all).
- **Explanations and reference added.**

p8319, l3f: "A CIR decrease over time might reduce the sensitivity" – Unclear. Does not make sense to me. CIR is an expression of the atmospheric state. E.g., thin clouds produce smaller CIR. Still, they do not effect the sensitivity (for a given atmsopheric state, the method is still sensitive to the same clouds).
- **The CIR decrease might be due to instrumental degradation which would lower the sensitivity. Added.**

p8319, l13: What are you referencing by Hommel et al.? You are discussing your own results here, aren't you?
- **Reference might be not adequate here, deleted.**

p8319, l17: "A considerable increase in CTH was found starting at the end of 2007" – How long did it persist? I can see it exclusively at turn of 2007/08.
- **It persisted from 12.2007 to 03.2008, corrected**

Fig. 7: "The MSG picture [. . .] has been adapted." – How? Why? Do you mean: 'Adapted from Reuter&Pfeifer'? Does (b) imply SCIAMACHY is looking straight south? is it like this? or is the image not geographic-north oriented?
- **Corrected. At the equator SCIAMACHY looks south, MIPAS north**

Fig. 8: Is it correct that CIR are 14–22? seems very high.

- **For visualization purposes it has been multiplied by 10. So it really is between 1.4 and 2.2.**
- **Note added**

Q4.100: Fig. 9: Can you clarify how height resolved coud fractions are defined? Such that sum over all layers gives total CF?

- **Correct.**
- **Explanation given in text.**

Generally, why are the comparisons only done or presented over limited periods of time instead the full lifetime of the instruments?

- **For Nadir comparisons we have used the period 2003-2009, for MIPAS 2008-2012.**
- **At least for nadir an update is possible, while for MIPAS the period seems best suited because of global data availability**.

p8320, l14: "as in limb only the highest CTHs are detected in a grid cell." – in my understanding nadir also detects only the "highest" CTH (or are SCIA nadir CTH multi-layer cloud heights?), just along a nadir LOS it takes longer (hence, deeper into the atmosphere) to reach the necessary path optical depth to get a radiometric effect.

- **True. But within a nadir scene the "limb clouds" would be those with the highest nadir CTH. Rephrased.**

p8320, l15: "Only data were used for the following SACURA quality checks:" – Sentence does not make sense. Parts missing?

- **Rephrased.**

p8320, l20f: "Very high limb CTHs over South America and South Africa were only partly seen in nadir data." – This is inconsistent with your Figures. CTH over South America and South Africa are higher in (modified) nadir data than in limb data!

- **The nadir CTHs were higher over the indian ocean. But adding a standard deviation can lead to overestimation in areas of high SD values.**

p8321, l5: "the nadir retrieval is restricted to the detection of water clouds" – Do you say that SCIA nadir can not detect any ice clouds?

- **Corrected and reference given**

p8321, l14f: "The black line and the blue area depict the SACURA nadir results and its 1-_ scatter." – No blue area in my Fig13.

- **corrected**

p8321, l26ff: "Better agreement can be achieved with other limb viewing instruments" – Better agreement maybe, but other limb measurements suffer from the same difficulties as SCIA limb. Hence, I'm sceptical about the independence of the two measurements and about the qualification of the MIPAS data as reference data set. Furthermore, with SCIA looking forward and MIPAS backward, one can expect the smearing/cloud displacement effects to act in different directions. Please discuss implications.

- **For a first validation, MIPAS is perfect and was also validated with CALIOP data, where MIPAS CTH were 1 km higher. Reference added.**

p8322, l8f: "we do not see the gaps in vertical direction of the SCIAMACHY cloud top heights." – Reformulate properly. Of course we see the gaps in the SCIA data. Why should we see SCIA gaps in MIPAS data?

- **Rephrased.**

p8322, l11f: "But there is [. . .] MIPAS top heights that clearly lie outside the vertical field of view" – Outside of which VFOV? MIPAS'?

- **Both. Rephrased**.

p8322, l15: "where the lowest possible MIPAS tangent heights were at about 10 km,

which can partly explain the differences." – How does this explain differences for individual collocated measurements? If the cloud is high enough for MIPAS to detect it (else it would not be in the scatter plot), SCIA sould be able to detect it, too, would it not?

- **Added explanation.**

Q4.110: p8322, l18ff: "The vertical differences are around −1.1 km [. . .]. This difference can be mainly attributed to the different tangent height step sizes" – I am not convinved. Fig15 shows a systematic tangent height dependency of the differences. Explain, how this is in line with the different vertical resolution you attribute it to. If it were an resolution isses, I'd expect random differences.

- **Added explanation. MIPAS had also an offset of 1km compared to CALIOP, reference added.**

p8322, l24ff: "Sembhi et al. (2012) compared MIPAS [. . .] and CALIOP [. . .] and found that MIPAS CTHs are in the range of up to 1 km higher for altitudes between 12 and 20 km, which is in line with our comparisons." – I find that a too far-fetched conclusion. Please compare SCIA limb CTH and CALIOP data directly.

- **For an extended validation of this CIR and the IR-CIR a comparison with CALIOP is planned. Beyond the scope of this paper.**

p8323, l5f: "Results from the lowest tangent height should be taken with caution, as MIPAS has latitude dependent tangent height cycle." – How does this affect MIPAS CTH? I would expect that MIPAS retrieved CTH are corrected for varying tangent height, and when comparing just collocated positive cloud hits (that is what you are doing, are you not?), one excludes cases with CTH too low to be seen by MIPAS. Saying that, it would be interesting to see a contingency table over all (cloudy and non-cloudy) collocations.

- **Paragraph has been updated**

p8323, l6ff: "Higher cloud top heights [. . .] at the highest SCIAMACHY tangent height 18.5 km. This can partially be explained by the coarser tangent height step size of SCIAMACHY, as the 18.5 km height is generally above the tropopause." – How does this explain the difference? Is MIPAS restricted to tropospheric cloud detection?

- **No, but has 2 measurements with oversampling instead of one from SCIAMACHY with undersampling. Added explanation.**

p8323, l27: "zonal mean CTH for clouds between 6 and 28 km" – Why globally from 6km if MIPAS has issues measuring CTH lower 9 or 10km in the tropics?

- **This value was used in Spang (2012).**

p8323, l27f: "Furthermore we have compared [. . .]" – Please show results. You keep claiming good agreement without allowing the reader to judge by oneself.

- **Right, added superimposed SCIA results to figure from Spang (2012).**

p8324, l2f: "the interhemispheric CTH differences in the SCIAMACHY data had disappeared." – Why is this here? No IH-differences discussed in this section before. How is that related to MIPAS data?

- **deleted**

p8324, l10f: "as the CIR _ is strongly height dependent. Because of the lower values of _ towards the surface the extra scattering due to aerosols have a high impact." – Neither of this has been shown before. Add in Sec3.

- **Added as figure 1 in supplement and updated figure 6**

p8324, l13: "But as shown in the cloud fraction distributions," – This has at best been speculated there, not shown (as in giving some evidence).

- **Evidence given by reference to CALIOP**

p8325, l12: "was estimated for both volcanoes from their measurements." – From whose/which measurements?

- **Anderson, referenced**

p8325, l23: “the lowest layer [. . .] is situated” – lowest layer of what? Inconsistent with layer range discussed above.

- **Tangent height, corrected**

Q4.120: p8325, l26: What is occurency frequency? Same as/related to cloud fraction? How has it been derived? Like CF in Sections before (e.g. 2x2_ cells)?

- **Only COF now used and defined**

p8326, l12f: “Overall about 45% of all measurements in the lower stratosphere” – measurement equals “cloud” hit?

- **corrected**

p8326, l20: “They detected the plume” – Who? Doeringer et al.? It reads odd to refer to subjects that have only been introduced in parenthesis-cites before.

- **OK, changed**.

p8327, l5ff: “Although the use of only one threshold for all atmospheric situations is simplistic in nature, it was shown that it was sufficient for the majority of cases” – I disagree. It was not demonstrated what happens, when threshold is changed (what do we know how many cases where not or falsely detected?), i.e. it can not be rated as “shown”.

- **Checks were made for a higher threshold 1.5 leading to similar results.**
- **Lower thresholds would lead to more false detection in the lower troposphere (see new Fig.6)-**

p8327, l13f: “SCIATRAN model studies have shown that the method is very sensitive for a wide range of cloud optical properties and cloud top heights.” – Not in this paper, though. At least I do not call two cloud cases with only cloud height varied “a wide range”.

- **Now 4 cases, which are chosen for one end of the range, which the lower one.**
- **I.e. Very low COT were simulated and moderate AOTs used in one case.**

p8327, l20f: “SCIAMACHY CTHs were generally lower by about 1 km, which can be explained by differences in the tangent height steps.” – Disagree. Not properly explained to readers.

- **More explanations added**.

p8328, l2: “It was shown that the height differences can partly be attributed to the different sensitivities of the viewing geometries.” – Was not shown, but stated.

- **rephrased**

p8328, l6f: “Unfortunately, SCIAMACHY was the only instrument so far that was capable of making use of the two viewing geometries.” – I do not see a strict requirement of having both realized within one instrument. Two simulataneously measuring nadir/downlooking and limb instruments should provide equally good data (maybe even better, since continuous data).

- **Right, refomulated**

p8328, l11: “The use other wavelength pairs in the near IR [. . .] enables to distinguish between clouds and aerosols.” – Please provide a reference proving that it is indeed (and not just might be) possible or reformulate the statement.

- **Reformulated**

3 Technical corrections

Throughout the paper: Use consistent spelling (e.g., Nadir vs. nadir), terminology (e.g., CIR or _ or CIR _, cloud fraction vs. cloud occurence frequency, Figure or Fig.) and tenses (e.g., p8310 “We have studied” vs. “colour index ratio will be studied”, p8310 “ENVISAT moves/flies” but “zenith angles occurred”, p8312 “VFOV was at” vs. “VFOV edges are then”, etc.). Unit specifications for plots belong in the figure captions (or the plots themselves), not in the text.

- **Corrected**

Q4.130: p8297, l25: “infra–red” ! infra-red
- **Corrected**

p8303, l2: “The full Earth was covered after 6 days at the equator” – Reformulate. Equator can never be full Earth.
- **Rephrased**

p8304, l22: Either “annual cycle” or “inter-annual variabilities”.
- **Corrected**

p8305, l4: & p8306, l18: “SACURA retrieves” and “The limb sounder MIPAS retrieved” – Odd language. What is the retrieving entity, the retrieval algorithm or the instrument?
- **The instrument retrieves light and the retrieval algorithm retrieves data.**
- **Rephrased**

p8305, l25ff: “Due to the specific spectral measurement range MIPAS [. . .] as a pure limb sounder recorded more limb scans [. . .]” – does not makes sense. reformulate.
- **reformulated**

p8310, l26: “the highest sun zenith angles” ! the largest sun zenith angles.
- **Corrected here and at other places**

p8311, l3: “SAAs larger than 90_” ! “SAAs around than 90_”?
- **corrected**

p8312, l13: “CIR was highest for very low SAA and lowest around 90_ SAA, which is in the tropical region.” – which of the two is (rather: occur) in the tropics? Besides, lowest CIR seems to occur for high SZA (south polar region).
- **corrected**

p8312, l18: “where the threshold was not reached (see Table 1).” – wrong reference? Table 1 does not deal with any thresholds.
- **corrected**

p8312, l27: “aerosols are within the first 5 km” – Reformulate properly (first of what? along LOS? from TOA? . . . ?).
- **corrected**

p8322, l22: “field of view increased” – wrong tense (at least, i expect it is always increasing).
- **corrected**

Q4.140: p8315, l3: “that the line-of-sight crosses” – rather: that the (V)FOV crosses/intersects.
- **rephrased**

p8317, l23: “High fractions were also detected at the height range (Fig. 9b).” – Parts of sentence missing?
- **Corrected**

p8318, l21: “a rather stable period from the start” – Lab slang (start of what?). Formulate properly.
- **Rephrased**

p8320, l9: “low geometrical limb resolution” ! low horizontal (or along-LOS) resolution of limb measurements.
- **Corrected.**

Q4.144: Figure captions: Reformulate captions with correct use of sub-figure references. For exmaple, “Global distribution of [. . .] (b) to illustrate [. . .]” and “Global annual mean [. . .] (b) the limb cloud fraction” are no proper sentences and do not make sense.

- **Done**.