Point-by-point response to the reviews:

We thank the reviewer for the helpful comments and suggestions. Based on the comments of both reviewers, the paper has been revised. In particular, section 2.2 gives now more details on the effect of potential aerosol detection, cloud inhomogeneities and includes already the introduction of the term surface area density path (ADP) instead of section 5; a couple of new subsections are incorporated in section 3 for comparisons to SAGE II and ground based lidars; a new section 5.7 on the formation process of LMS clouds in the CLaMS model is added.

Reply to Review#1:

A major concern is that a simple radiance color ratio i.e. the cloud index (CI) appears to be a fairly blunt tool for extracting this kind of detailed information on thin cirrus. Although, the CI works fine for general cloud detection there are limitations not addressed here. For example what is the discrimination between thin cirrus and aerosols? A single CI can only provide a threshold discrimination (regardless of particle type) although the expected distribution for CI from aerosols which almost certainly overlaps the CI distribution from thin cirrus is not discussed. How do we know that the LMS "bump" in the tropical CRISTA COFs is not stratospheric aerosol? The CRISTA CI COF are just compared directly with the CALIOP IWC COF. Potential differences between cloud detections using the CALIOP level 2 IWC (which I believe is screened for aerosols) and an analysis based on horizontally averaged level 1 backscatter is discussed by the authors only in the context of ultra thin cirrus.

Reply 1.1: The reviewers comment is correct. The importance of aerosol was not explained in sufficient detail in the old version of the manuscript and a differentiation between aerosols and ice particles was not explicitly applied to the observations. In response to the comment we added more details on the aerosol sensitivity, the tests we performed, and already published results related to the CRISTA measurement period in section 2.2:

Emissions by aerosol particles in the UTLS region can affect the CI values similar to clouds and a differentiation between ice and aerosol particles would be useful. Due to the very low volcanic activity after the major eruption of Pinatubo in 1991 for the rest of the 90s, the year of the CRISTA mission 1997 and the following years are a period of exceptional low UTLS aerosol load in the tropics (e.g. Vernier et al., 2011) and at mid-latitudes (SPARC, 2006). The long term records of measured aerosol extinctions by the satellite instruments SAGE II (at 0.5 and 1 μm) and HALOE (at 5 μm) for 1997 show extremely low values for this period (SPARC, 2006). Radiative transfer model (RTM) calculations with extinctions transferred from the HALOE wavelength to the 12 μm region of CRISTA as well as for particle size distribution from balloon measurements at mid-latitudes (Deshler et al., 2003) have been performed to quantify the effect on the cloud index approach. The results show the aerosol equivalent CI values in the range of 4 to 6. A CI<3 cannot be achieved for 1997 conditions (see also Spang et al., 2002). Consequently we can exclude that the detected events with $CI_{thres}=3$ in the following analyses are related by radiance emissions of the background aerosol or caused by an injection of volcanic aerosol in the tropics or mid-latitudes.

Further, for a more direct test we have applied to the data a new method for the differentiation of sulphuric acid aerosol and ice particles developed for the MIPAS instrument by Griessbach et al. (2013). The method uses additional wavelength regions at ~960 cm^{-1} and 1225 cm^{-1} to provide an aerosol/ice classification. The different spectral slope of the continuum like emissions for ice, H_2SO_4, and volcanic ash makes the differentiation of these potential aerosol and cloud types possible. If we apply the detection restriction of CI<3 the new method shows no aerosol but usually ice signals for most of the CRISTA spectra around the tropopause. This analysis corroborates the conclusion derived above on the basis of RTM calculations that values CI<3 cannot be caused by stratospheric aerosol.

We also added a comment on the potential source (wildfires) for the ‘tropical bump’ in the COF distribution in last paragraph of section 3.2. We kept this as a brief discussion because the tropics are not the focus of this study:

Exceptional events like wildfires with pyro-convection are potential sources for unexpected enhanced UTLS aerosol load (e.g. Fromm et al., 2010) and may cause the unexpected secondary maximum above tropical tropopause in the CRISTA COF results. High fire activity was observed in the time frame of the CRISTA mission and the following months over Borneo caused by the strong 1997/8 El Nino (e.g. Wooster et al. 2012). We checked the longitudinal distribution for the tropical CTH events above the tropopause, but found no specific clustering of these events above the Borneo region. In contrast, the observed regions are inline with the typical regions for high subvisual cirrus cloud occurrence in the tropical tropopause region (e.g. Wang et al., 1996). For summer conditions these are the warm tropical landmasses and the Micronesia area – the so-called warm pool region. Due to the specific El Nino conditions the latter region was extended in direction to the mid pacific (e.g. Spang et al., 2002) which is also observed in the CRISTA cloud distribution above the tropopause. The tropical secondary maximum is so far not observed by other sensors. Based on the data investigated here, we cannot decide whether the occurrence of subvisual cirrus well above the tropopause is a regular feature or a particular event in the time frame of the CRISTA mission.

The authors need to do more to convince the reader that 2-km vertical sample interval coupled with a 1.5 km field-of-view can yield reliable information on sub-kilometer thick clouds. No doubt that for an idealized unvarying unbroken train of cirrus the CTH from a limb sounder can be determined accurately. More convincing would be to take examples of contiguous CALIOP IWC profiles and run radiative transfer calculations for the CRISTA line of sight to simulate CRISTA CI values for a realistic "known" input.

Reply 1.2: Uncertainties in the CTH by thin layer clouds or clouds covering only the lower part of the vertical FOV are investigated in detail in section 3.3. Our simulations therefore cover the mot important uncertainties introduced by the vertical FOV and sampling for horizontally homogeneous cloud layers along the line of sight. Implementing ‘realistic’ cloud fields from CALIOP IWC retrievals, like suggested by the reviewer, may add additional information for single case studies. In addition, 3D- or 2D-inhomogeneities for clouds are very difficult to implement in radiative transfer models. A first attempt for a limb ray tracing through high resolution 2D IWC distribution from ECMWF analyses in combination with a forward model including multiple scattering effects is presented in Spang et al. (2012). Cloud detection results are compared with the original input IWC fields. The analyses confirm that the detection sensitivity is mainly related to the integrated surface area density path (ADP) of the cloud particles along the line of sight and the CI_{thres} can be attributed to a fixed detection sensitivity in ADP (see revised section 2.2). That study also shows that horizontal broken cloud fields do not need to be considered in detail in a model study. Due to the integrated view of the limb sounder significant differences in the retrieved cloud top height are not expected for two clouds with constant ADP. For example a large IWC concentrated over a short horizontal distance or even multiple cloud fragments and a low IWC distributed homogeneously over a large horizontal distance result in very similar CTHs. The study shows that size of the FOV and location of the CTH in the FOV are the driving factors for the CTH error.
Consequently, for modelling a large variety for ADP scenarios in a RTM study it is sufficient to vary the extinction at the tangent height layer (compliant to variety in IWP and IWC). This implies a homogenous cloud filling over the complete tangent height layer. This simplification allows us to use a much faster 1-D forward model without scattering and by this it is possible to consider a large variety of extinctions, geometries (CTHs) and atmospheric background conditions (tropical, mid, and high latitudes). The effect of the vertical FOV and sampling issues on the retrieved CTH are then applied to the modelled radiance spectra with high vertical sampling (for example 100m, see section 2.3).

We included additional information into and reworded the last paragraph of section 2.3:

Uncertainties in CTH determination from broken cloud segments along the line of sight in combination with the horizontal integration of the limb information and from the cross track extension *of the FOV (30 arcmin, ~15 km) are not considered in the present analysis. Spang et al. (2012) showed for the combination of realistic limb ray tracing through high resolved 2D IWC fields from ECMWF analyses and RTM calculation that due to the integrated view of the limb sounder significant differences in the retrieved cloud top height are not expected for two clouds with constant ADP. For example a large IWC concentrated over a short horizontal distance or even multiple cloud fragments and a low IWC distributed homogeneously over a large horizontal distance result in very similar vertical CI profiles and corresponding CTHs. The study of Spang et al. (2012) showed that the size of the FOV and the location of the CTH in the FOV are the driving factors for the CTH error. However, due to the limb geometry cloud Inhomogeneities along the line of sight can result in an underestimation of the CTH with respect to the true CTH. If the cloud segment is placed significantly in front or behind the tangent point (>150 km for a FOV of 1.5 km) and consequently at higher altitudes then the retrieved CTH will be negatively biased. This error would therefore not explain the cloud observation above the tropopause.*

Even better would be to do the test with level 1 CALIOP backscatter data since the COFs for the two instruments are so discrepant in the tropics. So why not download the CALIOP data and make the appropriate comparisons? This is clearly relevant to the scope of the present work and results could be plotted in Fig 4 (which lacks a CALIOP comparison in the two bottom figures) instead of picking up a couple of plots from another publication (which is for a much longer 3 month composite JJA time frame).

Reply 1.3: A direct comparison of the CRISTA and CALIPSO measurements is not possible because of the different period of the measurements. Of course a detailed analysis of the CALIOP level 1 data would be very interesting and relevant for the scientific questions raised in this paper. However, this needs to be a very careful and complex analysis of the dataset (for example improvements of the signal to noise issues in CALIOP). In addition it is not straight forward to compare limb and nadir information, like already discussed in the manuscript (section 3.4). The realisation of such a reanalysis of the CALIOP data would represent a self-contained study. Here we like to focus on the observations of LMS cirrus clouds by CRISTA and just a brief comparison with the current status of research in the literature like the CALIOP data analyses.
We added a small paragraph at the end of the summary section to highlight the importance of a reanalysis of the CALIPSO for more detailed and better information on LMS cirrus clouds:

As a next step, optimised analyses of the currently best suited spaceborne and still operating instrument for cloud studies, the CALIOP instrument on CALIPSO, would help to improve the knowledge on cloud cover and frequency distribution on temporal and spatial scales. Spatial averaging of high resolution CALIOP level 1 data have to be used to improve the detection sensitivity for optically very thin clouds.

The authors state that the detection sensitivity "is linked to the detection threshold and depends to some extent on the seasonal variation in the trace gas concentrations in the applied spectral windows" (P12311:L16-19). This is a key point. However, note that there are also potential large vertical variations of water vapor within a single profile occuring near the tropopause. In fact located in the tropics, just above the tropopause, is the region known as the hygropause (e.g. Teitlebaum et al, GRL, 27, 211, 2000). Since the water vapor may be as low as 2 ppmv in this region it would be necessary to know if such a local H2O minimum could generate an apparent increase in the CI and therefore lead to a bump in the cloud occurrence fraction. On P12330:L20 we have CI = I(788:796) / I(832:834). The water vapor line used is given as 12.7 um on P12330:L20 (or 787.4 cm-1 and maybe affecting the numerator wavelength region) and since the CI value would presumably decrease along with a decrease in water vapor (there are no strong H2O lines in the denominator wavenumber region) false positive cloud detections may ensue (i.e CI < CI_threshold). This issue can only be resolved by radiative transfer calculations.

Reply 1.4: The sensitivity of the cloud index method with respect to water vapour abundance has been investigated in detail by radiative transfer calculations in previous studies (Spang et al., 2005, 2007, 2012, Sembhi et al., 2012). In this studies the cloud-like effect of high tropospheric water vapour mixing ratios by the continuum emissions is analysed. The 12.7 µm emissions for H_2O mixing ratios < 10 ppm are very weak and only a small contributor to the mean radiances of the relatively broad wavelength region of the cloud index numerator (MW1). Therefore, so far, variations as in stratospheric water vapour were not spotted as an error source for the detection sensitivity or for a candidate for false detections. But the reviewer comment is plausible and the variation in CI due to the hygropause-like effect should be better quantified.

We made some new radiative transport calculations with a varying water vapour layer of 2 to 4 ppmv above the tropical tropopause between 16 and 18 km altitude. The setup is similar to the sensitivity tests in section 2.3. The change in the CI numerator for the 4 to 2 ppmv model runs causes a reduction in the radiances of ΔMW1 < 3 nW/cm^2 sr cm^{-1} or < 0.4% for the MW1(4 ppmv) signal. This gives a maximum changes in CI of <0.3% (or $\Delta CI_{max} = 0.02$). Such small systematic changes in CI have only a marginal effect on the detection sensitivity and cannot be responsible for the enhanced cloud occurrences above the tropical tropopause.

A short comment on this potential water vapour effect is now presented in section 2.2, end of the second paragraph:

Radiative transfer calculations show that water vapour variations in the tropical tropopause region have only a marginal effect on the numerator in the CI colour ratio and the corresponding CI-threshold approach (even a factor of two, like observed at the tropical hygropause compared to the higher stratospheric value causes only a $\Delta CI_{max} = 0.02$*). The water vapour line emissions are to week to compete with dominating* CO_2 *and* O_3 *in the 788-796 cm*$^{-1}$ *region.*

Minor comments:

All minor comments are incorporated into the manuscript, replies to more detailed comments are stated below:

P12328:L17 Better to state the horizontal extents and vertical thicknesses here. i.e. what is the minimum thickness of the 100 km cirrus cloud and what is the minimum extended length of the 1 km thick cirrus detectable for these IWCs?

In the introduction we intended to reference to quantities already presented in other publications. For a cloud layer of fixed size and not filling completely the vertical field of view (FOV) the detection limit becomes different for instruments with different FOV size (e.g. CRISTA and MIPAS). But the numbers requested by the reviewer are important and helpful for the reader and are now presented in section 3.4 (CRISTA-CALIPSO comparison).
Here we added also ‘vertically’ FOV to highlight that the number are not independent on the vertical extent of the cloud.
‘… the cloud has a vertical extent completely filling the vertical FOV’

P12328:L20 So what is the CALIOP IWC detection limit (for the current data product)?
We added this information at the end of the paragraph above.

P12340:L6 No significance testing is actually done here
Title of section 3.3 is changed to: *Statistical tests of …*

P12342:L24 This should be refered to as a backscatter channel. CALIOP only measures backscatter for which the units are km-1 sr-1. However, the units given here are km-1 as for an extinction. Or was a lidar ratio for cirrus applied to obtain the extinction values? For CALIOP the extinction is a derived quantity which requires some assumption for the Lidar ratio (extinction-to-backscatter ratio) which depends on the particle type.

The comment is correct, all presented extinction values assume ice particles. Details on the CALIOP extinction and IWC retrieval are given in Avery et al. (2012) and references therein. We changed the sentence to:

The current detection limit for cirrus clouds averaged horizontally to 5 km for the 532 nm backscatter channel results in a retrieved extinction value of 0.005 to 0.02 km^{-1} *(Avery et al., 2012), which represents an equivalent IWC of 0.1 to 4 x 10*$^{-3}$ *g/m*^{3}*.*

P12342:L26 I think you mean something like the IR limb sounder detection limit can be considerably better depending on the horizontal extent of the cloud. e.g. in the degenerate case a 1km long by 1km thick cloud would not practically yield a longer limb path than the nadir path.

Correct, we changed the sentence as suggested by the reviewer.

P12342:L27 Again, better to state the horizontal extents and vertical thicknesses here. i.e. what is the minimum thickness of the 100 km cirrus cloud and what is the minimum extended length of the 1 km thick cirrus detectable for these IWCs? Why is this quoted again here (see P12328:L17 above) and why are the units now in mg m-3?

We changed units consequently throughout the manuscript to g/m^{3} and present at this place more details on the equivalent IWC and IWP of the CRISTA detection thresholds compared to CALIOP, see also response to comment P12328:L17.

References:

Deshler, T., Hervig, M. E., Hofmann, D. J., Rosen, J. M., and Liley, J. B.: Thirty years of in situ stratospheric aerosol size distribution measurements from Laramie, Wyoming (41°N), using balloon-borne instruments, J. Geophys. Res., 108(D5), 4167, doi:10.1029/2002JD002514, 2003.

Fromm, M., Lindsey, D.T., Servranckx, R., Yue, G., Trickl, T., Sica, R., Doucet, P., and Godin-Beekmann, S.: The Untold Story of Pyrocumulonimbus. *Bull. Amer. Meteor. Soc.*, **91**, 1193–1209. doi: http://dx.doi.org/10.1175/2010BAMS3004.1, 2010.

Griessbach, S., Hoffmann, L., Spang, R., von Hobe, M., Müller, R., and Riese, M.: MIPAS Volcanic Sulfate Aerosol Observations of the Nabro Eruption, in: Stratospheric Sulfur and its Role in Climate (SSiRC), Atlanta, Georgia, USA, October 2013, SPARC, available at: http://www.sparc-ssirc.org/downloads/Griessbach.pdf, 2013.

SPARC, Assessment of Stratospheric Aerosol Properties (ASAP), Thomason, L. W. and T. Peter, Eds., SPARC Report No. 4, WCRP-124, WMO/TD, No. 1295, http://www.sparc-climate.org/publications/sparc-reports/sparc-report-no4/, 2006.

Wooster, M. J., Perry, G. L. W., and Zoumas, A.: Fire, drought and El Niño relationships on Borneo (Southeast Asia) in the pre-MODIS era (1980–2000), Biogeosciences, 9, 317-340, doi:10.5194/bg-9-317-2012, 2012.

Reply to Review#2:

Major point:

The formation of cirrus clouds in the lower most stratosphere (LMS) is not addressed in this contribution. However, it is very crucial if the clouds are formed in the troposphere and then are just advected into the stratosphere or if the clouds are directly formed in the stratosphere. The authors should use the model results to investigate this issue at least qualitatively. Although the model might not be able to reproduce the ice water contents in a correct way, the humidity values along the trajectories should give some hints about the formation and even on the location of the nucleation event in terms of stratosphere vs. troposphere.

We followed the suggestion of the reviewer and included an additional subsection 5.7 on *LMS cloud formation* in the CLaMS model in conjunction with trajectory calculations and CRISTA observatons. New Figures 11 and 12 are discussing the temporal evolution of air masses transported from the subtropical jet over the Atlantic and finally forming ice clouds at high latitudes.
Finally we added following paragraph in the summary section 6 on the results of this trajectory study:

Trajectory studies with the cirrus module of CLaMS in comparison with the CLaMS simulations show the importance of mixing for the formation of ice clouds in the LMS. Mixing events at the subtropical jet can generate the entrance of enhanced vapour mixing ratios into the LMS which favour the formation of ice clouds even at high latitudes over Scandinavia in the model in agreement with CRISTA observations.

Minor points:

1. Comparison with former investigations: I miss some former investigations, which might be used for comparison. Actually, there are some recent activities of cirrus climatologies using LIDAR data in the French community (Dionisi et al., 2013; Hoareau et al., 2012, 2013; Dupont et al., 2010, 2011), which might be used for comparison. The SAGE data analysed by Wang et al. (1996) are available as gridded data; thus, they can be easily compared with the CRISTA data in a climatological sense. Finally, Spichtinger et al. (2003) reported ice-supersaturated layers in the stratosphere. Since thin cirrus are often embedded into a supersaturated environment, these data might also be used for comparison.

We followed the reviewers suggestion and included a couple of sub-sections in section 3 on comparisons with SAGE II, specific humidity measurements, and ground based lidars (3.4, 3.5, 3.6). However, the comparisons are not quite quantitative, because most of the analyses are suffering under the lack of an exact tropopause determination and a reanalysis of the datasets would be desirable.

2. Broken cloud fields: The authors mention in the text that CRISTA is very sensitive to homogeneous thin clouds but most cirrus clouds are inhomogeneous. They mention some errors for broken cloud fields; however, these sources of uncertainties should be explained in more details (maybe in the appendix).

We like to refer referee#2 to our Reply 1.2 for referee#1. There we give more details on the effect of broken clouds. In addition, we included additional information into and reworded the last paragraph of section 2.3:

Uncertainties in CTH determination from broken cloud segments along the line of sight in combination with the horizontal integration of the limb information and from the cross track extension of the FOV (30 arcmin, ~15 km) are not considered in the present analysis. Spang et al. (2012) showed for the combination of realistic limb ray tracing through high resolved 2D IWC fields from ECMWF analyses and RTM calculation that due to the integrated view of the limb sounder significant differences in the retrieved cloud top height are not expected for two clouds with constant ADP. For example a large IWC concentrated over a short horizontal distance or even multiple cloud fragments and a low IWC distributed homogeneously over a large horizontal distance result in very similar vertical CI profiles and corresponding CTHs. The study of Spang et al. (2012) showed that the size of *the FOV and the location of the CTH in the FOV are the driving factors for the CTH error. However, due to the limb geometry cloud Inhomogeneities along the line of sight can result in an underestimation of the CTH with respect to the true CTH. If the cloud segment is placed significantly in front or behind the tangent point (>150 km for a FOV of 1.5 km) and consequently at higher altitudes then the retrieved CTH will be negatively biased. This error would therefore not explain the cloud observation above the tropopause.*

3. Model parameterisation: It is obvious that the model has problems in reproducing thin cirrus clouds. From my point of view, the very simple cirrus scheme suffers from the fact that as soon as ice is formed (at a given threshold of 100-150%) all excess water vapour is transferred to ice. This is not realistic, since the relaxation time is usually quite large depending on the surface area of ice crystals. …

The fact, that all water vapour above the threshold is transferred to ice can indeed lead to the formation of relatively thick cirrus clouds. This effect can quite clearly be seen in the new Figure 12 (bottom panel) of the revised manuscript. While the cirrus clouds in the model case with conventional freeze-out at 100 % evolve slowly over several hours, the cirrus clouds with freeze-out at a certain over-saturation are forming quite rapidly. Furthermore, conventional freeze-out leads in general to more and thinner ice clouds, since they form at higher temperatures. Nevertheless, as for the examples in Figure 12, the results concerning the evolution of the clouds once they have formed are quite comparable between these two cases.

An introduction of a relaxation time on the order of minutes or hours into the global model with its 24 hour time step would not lead to a significant change in the results. Regarding the trajectory runs with a significantly higher temporal resolution this would of course lead to thinner ice clouds, and in the following maybe to a higher potential for the generation of ice clouds afterwards, since there is a chance of less water being removed by sedimentation.

We are currently working on an implementation of higher temporal resolution for the cirrus parameterization into the global model, thus improving the results on the larger scale. In this implementation we will consider relaxation times on scales that we will try to deduce from results with a more detailed microsphysical bulk model (Spichtinger and Gierens, 2009a).

Technical comments:

Figures 1, 4, 5 are heavily overloaded and are hard to read. Please make them much larger and think about reducing the amount of information. Maybe it would be better to produce two figures instead of only one.

In the final version of the manuscript these figures will be significantly enlarged (confirmed by the ACP editorial office). We also reduced the information (wind contours) and changed the symbols in Fig. 5 to reduce the overload on different information.

References:

Spichtinger, P. and Gierens, K. M.: Modelling of cirrus clouds – Part 1a: Model description and validation, Atmos. Chem. Phys., 9, 685-706, doi:10.5194/acp-9-685-2009, 2009a.