Reply to Referee#3:

We thank the reviewer for the helpful for comments. Please find below the point-by-point response.

First, this paper has a tremendous amount of technical detail (albeit that is relevant and useful), and a substantial cross-validation of the cloud detection approaches and cloud frequency distributions with other sensors, but this manuscript is lacking new scientific insight. This paper appears to be more appropriate for Atmospheric Measurement Techniques, rather than Atmospheric Chemistry and Physics.

Based on the comments of both reviewers, the paper has been substantially restructured and revised. In particular, the discussion of technical details has been reduced (with a lot of material being shifted into the electronic supplement), while the discussion of new scientific insights has been extended.

Second, although the manuscript is not poorly written, it is unclear at times, and the flow between different sub-sections is at times poor. There are some occurrences of colloquial language and jargon terms that could be modified.

Following the reviewers suggestions the order of the result section 5 is changed, parts of the blind test section are rewritten (Sect. 4), and a few additional sub-sections are introduced for a better focus on the main ideas of the manuscript.

Third, and similar to the second point, it is frequently unclear why a given subject is mentioned or a figure is shown (20 in total, and most have multiple panels). There is a lot of information to process, and perhaps much of it could be condensed into references or in a more concise manner. The Appendices are a difficult slog in addition to the main body of the paper.

The amount of technical information is reduced in the revised manuscript by moving parts of the algorithm sections which are not crucial for the following analyses (altitude correction, Appendix 2 und 3) into the electronic supplement.

Specific comments:

p. 33015, lines 16-19: This is pretty fluffy language. Arguably, researchers have been working with high spectral resolution data more for than a decade (aircraft and satellite observations.
We changed the wording, included the 10 years of investigations, and mentioned the restriction to IR measurements in the limb.

p. 33016, lines 12-13: '. . .future ozone loss.' ?
Changed

p. 33018, line 7: '. . .either water or cirrus clouds.'
Changed

p. 33019, line 11: What is an 'FR-mode'? An example of technical jargon.
We tried to reduce the number of technical acronym and replaced this term with 'full spectral resolution mode'.

p. 33023, lines 5-12: Is this seasonal and latitude-varying climatology sufficient? If the trace gases (including water vapor) and temperature vary significantly, doesn't this impart important biases in the calculated radiances? Can't these variations be on the order of several Kelvins, and this can lead to missed/false cloud detections, and biases cloud retrievals? Perhaps this is not an issue because of some well-justified reasons, and the reviewer missed something along the way, but this would be a good example of a lack of clarity on this issue.

We added information and references on the topic temperature and water vapour sensitivity in the corresponding section of the new manuscript:

“Spang et al. (2004) already showed that the temperature dependence of the ratio is especially weak (e.g. <1%/K between 10 and 30 km for CI_A), but in the lower troposphere high water vapour amount (e.g >1000 ppmv for CI_A) can cause false detection results. Consequently, realistic water vapour variability must be taken into account when modelling GI values.“

“… and (c) UTLS water vapour variability is represented by a range of water vapour concentration profiles calculated from the saturation mixing ratio profiles with climatological temperature and pressure profiles (Remedios et al, 2007). In addition, per latitude band and altitude step the maximum H_2O from ECMWF re-analyses is used for a realistic representation of the large water variability in the modelled colour ratios (Sembhi et al., 2012).

p. 33023, line 13: The index here says “GI” (for gas index), but it says “CI” (for cloud index) on figure 3. Which one is it supposed to be?

It should be ‘Gas Index’ for this detection method. Text, figure caption, and figure were corrected.

p. 33028, line 5: ‘extinction’

Corrected

p. 33028, Section 3.2.2. This is another point of confusion. Earlier in the cloud detection methodology, the cloud top temperature and height were obtained (see lines 12-15 on p. 33020). Why is an additional retrieval of these properties needed at this step? Were the original CTT and CTH initial guesses to the actual retrieval of these properties, or did I miss something? Some clarity in the detection and retrieval flow is definitely needed here.

The sentence on CTT and CTP is a repetition. The original comment on p. 33020 is now deleted, because it was a confusing order shifted forward. We followed the reviewer suggestion and added a few sentences on the detection, estimated CTT and CTP and the macro-retrieval at the end of section 3.1. In addition we introduced for better clarity the combined summary parameter SUM_CLOUD for CTH, CTP and CTT already here instead of section 4.3.1.

p. 33028, line 19: The effective radius is not listed here. Is this retrieved after this retrieval step, separate from the CTH and CTT? This is quite confusing, although the flow chart of Fig. 1 suggests it is retrieved two steps later.

Effective radius is a separate retrieval step, because the applied algorithm needs the cloud type classification beforehand (like illustrated in Fig. 1). The simplified methods of the macro retrieval method with its applied microwindows are not sensitive to changes in R_{eff}. Consequently R_{eff} is not mentioned in this section.

p. 33030, lines 17-18: How about ‘A more detailed description of the classification is found in Appendix B and C.’

Changed accordingly

p. 33031, lines 6-7: It is never mentioned why only the top three cloudy MIPAS spectra have effective radius retrievals, and the lower spectra do not.

Explanation is now given in section 3.5:

“The current processor retrieves R_{eff} and ADP for the top three altitudes of cloudy spectra of the MIPAS measurement profile. This restriction is applied because tangent heights below are difficult to penetrate. A quantitative separation of cloud effects in the actual tangent height spectrum from emissions of cloudy layers above is not achievable with the current approach.”

p. 33031, lines 9-10: Are the authors suggesting that they can retrieve liquid water cloud properties from MIPAS? The reviewer is not aware of any limb emission or transmission study that has justified that this viewing geometry is able to determine anything about low altitude liquid water clouds. Please clarify this comment.

The information on liquid water cloud parameter is so far very limited. The main focus of the processor was to differentiate liquid water from ice particles, because the methods applied for microphysical parameter are dependent on the cloud type. Volume density path for spectra classified for liquid clouds were computed in the processor for exploration purposes and are not consolidated nor part of the current validation activities. Consequently we excluded now any misleading comments regarding liquid cloud parameters in the manuscript.

p. 33032, line 23: 'a significantly more compact'
Changed accordingly

p. 33033, line 24: 'However, the currently. . .'
Changed accordingly

p. 33033, line 24 to p. 33034, line 4: How do the authors know that this is a low bias? Have they made comparisons with other retrievals of ice cloud effective radius from other instruments? Or is this simply a guess given some initial results? It is true that the three reasons listed are all possibilities for a low bias, but did the authors actually quantify whether any of this is true, or is this simply speculation on their part? What about the fact that MIPAS effective radius retrievals are only obtained from the uppermost three layers? If effective radius is stratified with height (and this is suggested by in situ observations of cirrus clouds), couldn't this also be a factor, too?

How the low bias was quantified by the blind test retrieval (BTR) method is now explained in more detail in this subsection (the method is originally introduced in detail in the following section.)

"Modelled IR spectra based on defined particle size distributions, optical parameter, and radiative transfer calculations were fed into the processor for blind test retrievals (for details of the setup of the blind test retrieval see section 4.2). First results of the comparison show a significant low bias in the retrieved R_{eff} values compared to the input parameter."

The three specific reasons listed for the bias are not a speculation, because they are highlighting weak points come along when the R_{eff}-retrieval was derived from the database. The variation of R_{eff} with altitude is to some extent included in the size distributions estimated from the ECMWF data. The retrieval should be sensitive to these variations. But we agree with the referee, not all potential error sources have been investigated, further analyses and validation activities are necessary. A comment on the limited error analysis is now included in the paragraph.

p. 33034, line 6: What does 'simulated cloudy radiances for well-known cloud conditions' mean?
The term is changed to:
"A dataset of simulated cloudy and non-cloudy radiances for defined cloud geometries and parameters as well as microphysical properties was created for the validation of the processor by the blind test retrieval (BTR) approach."

p. 33034, line 11: 'and is based'
Changed accordingly

p. 33035, line 16: Liquid water clouds are mentioned again here. How are they detected in the initial stages? Are they assumed to be liquid based on the temperature of the cloud/MIPAS tangent point? How is a liquid/ice mixture handled in the radiative transfer simulations? Are there mixtures along the line of sight, in the vertical – how are they handled?

The detection methods of clouds in MIPAS FOV are independent from the cloud type (see also reply to p.33038 below). The handling of mixed and liquid clouds for the blind test is now described in more detail, e.g.:

"All scenarios (transects) included also areas where ice and liquid water clouds are present in the same ECMWF grid box (usually at altitudes below 5 km). "

And later on in the same paragraph:

"The sum optical properties for liquid and ice particle (extinction coefficients, phase function, etc.) are computed based on the sampling of the ECMWF grid. These optical properties are input to the radiative transport model, allowing mixtures of liquid and ice at a given point along the line of sight as well as the integration or weighted means along the line of sight."

p. 33038, lines 24-25: What does '. . .and it also has a tendency to. . .'
Corrected

p. 33039, line 1: Once again, liquid water clouds make it into the discussion. Are these assumed to be ice when they are retrieved? This is very confusing.

The detection algorithms are independent from the cloud phase. In Fig. 8 any cloud detection by the different methods should be confirmed by area density path ADP>0 for ice clouds (or volume density path VDP>0 for liquid water clouds, not shown). These ADP/VDP are computed straight forward from the IWC/LWC and effective radius information, which are used to create the modelled radiance spectra and represent the 'true state' or reference state for the blind test validation. At this point of the manuscript we like to highlight that the low cloud top heights in areas with ADP=0 are not artefacts but are caused by liquid particles in this specific regions. The text has been changed accordingly.

p. 33040, last paragraph of Sect. 4.3.2: The values of IWC discussed by the authors could be compared with the range of in situ aircraft values reported in the literature.
A comparison is now part of the manuscript:
"These values are in the typical range of IWC in-situ measurements in subvisual cirrus clouds. For example McFarquhar et al. [2000] measured IWC values from 10^{-4} g/m^3 - 10^{-6} g/m^3, similar to Lawson et al. [2008] with values between $5 \cdot 10^{-4}$ g/m^3 - 10^{-6} g/m^3. Davis et al. [2010] found in laminar SVC structures IWC values of significantly less than 10^{-5} g/m^3 over a vertical extent of a few hundred meters of the cloud. The authors showed that CALIPSO would be missing ~2/3 of this type of ultrathin cirrus clouds. "

p. 33041, line 9: What is a 'mean median'?
The misleading term has changed to 'median of the relative differences of ~30%'

p. 33050, line 4: 'point to a hypersensitivity'
Changed accordingly

p. 33050, 12-13: 'consistent, constant and small' is unclear
Paragraph has been changed to:

"However, at lower altitudes MIPAS seems to lose sensitivity with respect to SAGE II. This is in contrast to the GLAS comparison where MIPAS shows nearly constant mean differences (< 2 km, depending on detection method) at the lowest 2-3 altitude levels (troposphere). Increasing positive differences with respect to GLAS for the highest levels (polar stratosphere and UTLS) indicate a hypersensitivity of only 1-2 methods in some latitude regions but may point to slightly higher detection sensitivity than GLAS for ultrathin clouds. This was already postulated in the blind test retrieval results (Sect. 4.3.2)."

p. 33051, line 11: missing parentheses

Text has changed to: '… where the cloud top is located in the field of view from the tangent height ($h_{io} = TH – FOV/2$, …)

p. 33052, line 21: 'of the SVC flag'

Changed accordingly

p. 33053, lines 8-10: How does an observation of cloud around 20 km imply that watervapor of sufficient amount is being entrained into the stratosphere, and presumably clouds form there in situ? This argument doesn't make sense and the authors need to support this with some evidence. Couldn't this be a vertical smoothing effect because the weighting function in the vertical is nominally 3 km? Or could it be that overshooting convective towers are directly injecting ice particles into the lower stratosphere?

We agree with the referee that the argument is not evidential from the presented results. The data cannot really proof the argument of a stratospheric entrainment, consequently this term has been deleted. Vertically smoothing or uncertainties in CTH are of course a potential source of high cloud artefacts and is now mentioned in the manuscript:

"The observation of clouds above 20 km seems quite high. This is clearly in the stratosphere and due to the increasing temperatures and low water vapour mixing ratios (typically 2-4 ppmv) the formation of cirrus clouds becomes very unlikely. So far, it is not found in other remote sensing or in situ datasets. The cloud events in the stratosphere are potentially artefacts of the macro retrieval related to vertical smoothing effects caused by the coarse FOV and sampling grid (3 km). Over-sampling used in the MIPAS measurement since the end of 2004 (2 km steps) might help to reduce this effect."

p. 33054, lines 26-27: Fig. 19 seems to show that ISCCP has slightly more high clouds in the tropics than ATSR given the color scale.

We agree on this point (e.g. in the warm pool region) and changed the text accordingly.

p. 33054, line 27: You need to cite which AIRS product is being shown. Is this the AIRS NASA team, or the Stubenrauch et al. product?

A reference to the AIRS-LMD product is now presented (Stubenrauch et al., 2008).

p. 33056, line 15: What in particular is innovative with this confidence flag? MODIS has been taking this approach for years.

We see the point with respect to the nadir community/instruments, changed 'innovative' to 'new', and restricted the argument to the cloud detection in IR limb measurements.

Figure 1. Why are there separate boxes for 'detection' and 'cloudy'? The 'cloudy' box is for the assessment of confidence only?

Not only, the box is highlighting the most important decision step in the processing. A combination of plausibility test and the confidence flags are analysed to decide with high confidence about cloudy and non-cloudy MIPAS spectra. The diamond symbol is usually used for decisions in software flow charts.

Figure 2. Why does strong vertical striping in the field appear?

We made a note on this feature in the figure caption. It is caused by the quite accurate pointing to the fixed measurement grid (here 3 km) for MIPAS. The variability is getting larger for extended latitude bins, due a slightly drift in the nominal tangent heights with latitude.

Figure 7. Get rid of blue background – make it white. Also, the vertical axis should be changed to 0-20 km to be consistent with Fig. 8. Also, what is going on with the large amount of IWC below 5 km in the tropics? Is this really liquid water cloud and it is being called ice by the algorithm? This comment also dovetails with the previous comments about the lack of clarity on how liquid water clouds are handled in this algorithm.

The figures are changed accordingly and present the originally IWC of the ECMWF model only. In addition, corresponding LWC distributions for the liquid phase of water are used for the input of radiative transport model (described in section 4.2, but not shown). It is out of scope of the paper to decide if large amount of IWC below 5 km of the ECMWF model is realistic or not. We used the ‘model world’ to define a ‘realistic’ cloudy atmosphere for altitudes above 3km. MIPAS measured only down to 5-6 km in the tropics.

Figure 12. The lines should be distinguishable from each other. Why not use color?

We changed this with respect to the other colour coded figures with various detection methods.