Final author comment

The authors thank the reviewers for their helpful comments. In the following reply we have inserted the original comment in *italic face*. The replies are printed in normal face.

Anonymous Referee #1

Comment: *General Comments: This manuscript describes the retrieval of MIPAS HCFC-22 profiles from 2005-2012, validates them, and presents some figures showing its morphology. It concludes with a discussion of stratospheric trends from this short data set. It is long and poorly written, especially the retrieval section.*

Reply: The entire paper has been reworked with respect to language. Our reply to the criticism of the retrieval section is found under the respective specific comment.

Comment: *The validation section underused (maybe misused) the ACE data, which are a valuable tool for the validation.*

Reply: Misuse of data, we understand, is if we draw conclusions from the data which are not supported by the data. We could not identify any misuse in that sense. We comment on "underuse" in the context of the specific comments.

Comment: *Two balloon profile data sets are also used but are not appropriate for validating the MIPAS data.*

Reply: We disagree and comment on this in the context of the specific comments.

Comment: *There is little analysis of the MIPAS data. Figures are presented and described but not analyzed in any quantitative way. There is no HCFC-22 climatology in the Climatology section.*

Reply: There is quantitative analysis with respect to data quality assessment, growth rates, and time series. In the revised version we avoid the word climatology, because it causes misunderstandings and causes wrong expectations. Further, the dichotomy between qualitative and quantitative does not exist. Morphologies as discussed in Section 5.1, while not fully quantitative, are far more than only qualitative (R. Carnap, Introduction to the Philosophy of Science, Dover Publications, 1995).

Comment: *No meteorological data are used to support descriptions of HCFC-22 behavior or statements about processes affecting the distributions.*

Reply: We consider the use of meteorological data as questionable in the context of 2D distributions.

Comment: *Many statements about HCFC-22 stratospheric structures are known from previous studies (e.g., using ACE satellite data).*

Reply: While there exist trend studies for HCFC-22 (Rinsland et al, 2005; Brown et al., 2011) and budget studies for chlorine and fluorine from ACE-FTS (Nassar 2006a,b; Brown et al., 2013, 2014), no studies of global HCFC-22 distributions from ACE have been published, except those referenced in our paper. These publications do not contain the features discussed in our paper.

Comment: *The topic and the MIPAS data set used do have a place in the literature, so there is good 'raw material' here for a publication. However, this manuscript requires major revision. Some sections can be eliminated, others combined, and almost all need reorganization or rewriting. The comments below are organized by section; they note what I think the major problems are and how each section might be improved. Because of the significant revisions needed, it is premature to make many minor or technical suggestions.*

Reply: We comment on this in the context of the specific comments.

Comment: *The standard of English used is fair. I appreciate how challenging it is to write a scientific article in a second language! I strongly recommend having a native English speaker proofread the manuscript to improve readability. As an example 'The profiles obtained are' sounds better than 'Obtained profiles are'. A satellite measurement is a sounding but measurements aren't 'sounded. Also, the writing style uses many unnecessary words. E.g., 'has a more or less seasonal cycle. more or less adds nothing, try simply 'has a seasonal cycle'. 'So-called level 1B product': why 'so-called'?*

Reply: We have tried to improve the language of the manuscript and the paper has been proofread by native English speakers. However, language editing is now routinely done by professional language editors of ACP as part of the final publishing process. Thus, such issues should not affect the acceptance of the paper. In the manuscript we have never written that measurements are sounded. We have not written 'a more or less seasonal cycle' but 'for all altitude/latitude bins, a more or less pronounced seasonal cycle is...', which says that the seasonal cycle is more pronounced for some and less pronounced for som other bins. 'So-called' because 'level 1B product' is not a generic term but internal technical language of a specific community.

Comment: *Introduction*
p. 14786, Line 10. The most current determination of the atmospheric lifetime of HCFC-22 can be found in the 2013 SPARC report (a WMO publication). The SPARC lifetime is the same as reported here but it's a more current assessment of the lifetime.

Reply: The reference has been replaced.

Comment: *Section 2, MIPAS data*
No mention is made of the MIPAS sampling pattern. What is its latitude range? Is the range covered daily? Does is sample more at some latitudes than others? Does it measure all latitudes in all seasons?

Reply: The vertical sampling pattern is described in the original submission at the end of Section 2. The missing information on the horizontal sampling has been included in the revised version of the manuscript.

Comment: *Section 3, Retrievals*
This section is not well organized and it reads like a series of unconnected details regarding the retrievals.

Reply: We do not understand the comment. In the first sentence, we state in general terms what type of analysis method is used. Such an analysis method connects the retrieval space with the measurement space, using some constraint and prior information. This formalism is not repeated here because this would be redundant with many other publications, e.g. von Clarmann (2003) which is referenced at the end of the preceding section. In the following sentences we specify the retrieval space, the measurement space, and the constraint including the prior information. The second sentence specifies the measurement space in a sense that all tangent altitude are analyzed in one single step. The third sentence specifies the retrieval space (what are the variables; how are they discretized). In the fourth and sixth sentence we justify the use of the constraint and specify the constraint. The remaining part of this section discusses in more detail the spectral windows used and the problem of interference by other gases implied by the spectral analysis windows. It is not clear to us why this (admittedly but intentionally very short and compact) specification of the general method described elsewhere at length should be "a series of unconnected details".

Comment: *For example, the last paragraph of the section is on information displacement: how is this relevant to understanding the data?*

Reply: Information displacement is relevant in the same sense as the usual averaging kernels except that the information displacement and information smearing is connected to effects in the horizontal, not in the vertical. Thus we think that this is essential part of the data characterization, as are the vertical averaging kernels.

Comment: *Stiller et al, ACP 2008, on SF6, provides a good example of a retrieval description. I suggest you consider this and other MIPAS publications to see how they explain their retrievals (e.g.,) – then rewrite accordingly.*

Reply: The related Section in Stiller et al. 2008 contains a discussion of specific retrieval problems of SF_6 which had not been discussed before. These problems needed a thorough discussion, thus their retrieval section is four times longer than that criticized here. In the case of HCFC-22 no such specific problems have been encountered, and the retrieval strategy closely follows our standard approach which has been described in about 50 (!) preceding papers. Thus we consider it justified to mention only the specific choices of processing parameters relevant to the retrieval of HCFC-22. It seems to us to be adequate to best possibly avoid redundancies, to keep related sections short, and to focus on the essental information.

Comment: *p. 14789, Line 20. What is a 'zero-level profile'?*

Reply: We use a profile which is constantly zero for all altitudes. Due to the particular regularization scheme chosen the choice of the particular vmr does not affect the vmr, we could as well have chosen 42 without any change in the results. The only thing that matters is that the a priori profile is chosen altitude constant. This along with the regularization matrix has the effect that the retrieved profile is a smoothed version of the true profile. The related statement in the paper has been reworded for clarity.

Comment: *How sensitive are the results to the a priori?*

Reply: This information is provided by the averaging kernels, which are included in the paper. The integrals over the averaging kernels are, at good accuracy, unity, thus our choice of the constraint does not imply any bias. This information has been added to the paper.

Comment: *p. 14790, lines 3-15. The writing does not flow well.*

Reply: This paragraph has been rewritten.

Comment: *Where do the temperature analyses come from? (state it and reference it)*

Reply: It had been stated already in the original manuscript (p14790 l7) that MIPAS temperatures had been used. A reference to von Clarmann et al. (2003, 2009) has been added in the revised version.

Comment: *line 23. 'Its dominating components' maybe 'primary components'. 5 things cant all dominate.*

Reply: We do not understand why one group of errors cannot dominate over another group. And if we apply this level of linguistic rigorosity, then 5 things can't all be primary either. Nevertheless, we have changed to 'most important'.

Comment: *p. 14791, line 2. What is regularization? It hasn't been defined.*

Reply: Although 'regularization' is a standard expression in inverse theory, this term has been defined in the revised version. Regularization makes a singular or close-to-singular matrix regular by application of a constraint.

Comment: *Section 4, Validation*
p. 14792, line 11. What do you mean by 'unsolved problems cause by different altitude resolutions'? All the validation comparisons you make involve measurements with altitude resolutions different than MIPAS, so this statement is puzzling.

Reply: Aircraft measurements provide measurements at one altitude level but usually no vertical profiles. To apply the averaging kernel profiles of the better resolved measurements, profiles are needed. All the other instruments provide profile measurements, thus the related problem could be solved (Either by application of averaging profiles or by concluding that the actual profile is smooth enough that this step is not needed). This explanation has now been included in the revised version of the paper.

Comment: *Lines 20-22. How does MIPAS sampling (latitude and season) compare to ACE?*

Reply: Information on the MIPAS sampling has been added in the MIPAS section. Some further information on the ACE-FTS sampling has been included here. However, since we compare only collocated measurements, not mean profiles, sampling errors are not an issue in the comparison.

Comment: *Section 4.1, ACE comparison (figures 4-7).*
By examining a limited latitude region and sorting by season, Figure 5 provides the most useful comparison of this section. The two data sets agree within their uncertainties below 20 km all the time while seasonal differences are revealed above 20 km. Why, then, do you combine all latitudes and seasons in the other

figures? So much information is lost.

Reply: We do not agree that the profiles really agree. The error bars of the profiles represent the typical uncertainties of a single measurement, estimated as the mean error over the sample. This information has now been included in the figure caption. They are important information in their own right but to judge if a bias is significant, the standard error of the mean difference is needed, which is approximated by the standard deviation divided by the number of profiles. These are shown as error bars in the middle panel of Figure 4 and are hardly visible. This is, because due to the large number of profiles available, the standard errors are very small. Thus nearly all of the biases are significant. This does not mean that the measurements are bad but just that even very small biases can become statistically significant if the comparison ensemble is large enough. In the revised version, the missing information has been added in the figure caption in order to avoid such misunderstanding.

We usually do all analyses separately for different latitudes and seasons, but often the characteristics of the differences turn out to be homogeneous. In this case we present global figures to save space. We are reluctant to overload the paper with redundant figures but if the editor requires it, we could provide additional figures as a supplement.

Comment: *In fact, Figure 6 gives the opposite impression as Figure 5: it shows continuous, nearly 1:1 agreement below 25 km or so. (Thus I cannot understand the statement that the points fall into 2 clusters, p. 14794, line 2).*

Reply: We think that it is exactly correct what we write: The red points are on average above the line and the blue and green points are on average below. Thus, we do not quite understand the concern of the reviewer.

Comment: *In general I find that the words written in this section do not align with what is shown in the figures. For example on p. 14793, line 11, The bias is significant at all altitude levels. It clearly isnt!*

Reply: The bias is larger than the standard error of the mean difference, thus we conclude that the bias is significant. Unfortunately we have missed in the original submission to mention that the standard errors of the mean differences are the tiny, hardly visible, error bars of the middle panel of Fig. 4. Figure captions have been improved to avoid this kind of misunderstanding.

Comment: *In Figure 7, there is a bimodal shape in the lower right histogram, but as global data are combined in this figure,...*

Reply: We routinely do this sort of analysis (and most of the others) for each latitude band and season separately. However, in cases where these additional figures do provide only redundant information, we prefer to publish only the global plots, in order to save space.

Comment: *... who knows why this shape occurs? (But a different comparison would probably reveal the answer.)*

Reply: We know why. In this particularly case the bimodal distribution simply represents polar (lower mode) and midlatitudinal (higher mode) measurements. The number of tropical collocations is small and has thus minor impact on the histogram. The ACE-FTS modes are clearer distinguished because the South-North component of the line of sight of MIPAS, which observes roughly in the orbit plane, is more pronounced than that of ACE-FTS, whose line of sight is directed towards the sun. Since the mixing ratio gradients in North-South direction are typically larger than those in East-West direction, the MIPAS modes are smeared more than those of ACE-FTS. This information has been added to the text.

Comment: *Overall, the ACE comparison is inadequate. It could be a more valuable and useful part of this paper with comparisons that examine specific seasons and latitude ranges.*

Reply: We think we have refuted the criticism of our MIPAS-ACE comparison above.

Comment: *Section 4.2, Cryosamplers*
p. 14795, lines 3-6. These sentences are poorly worded. They serve as the introduction to the topic of the next paragraph and should be combined with it.

Reply: These sentences have been reworded and this and the following paragraph have been combined.

Comment: *The coincidence criteria is so broad (1000 km and 24 h) that when the profiles do not agree you really can't know why. If you use some meteorological analyses to show the profiles are from similar environment, then you would know whether it made sense to compare them.*

Reply: We disagree. We have plotted all MIPAS profiles matching the coincidence criteria around the balloon geolocation, and from their spread we can estimate which part of the difference can be attributed to natural variability. Furthermore, we have made a similar comparison for CFC-11 and CFC-12, involving the same set of measurements (Eckert et al., 2015, AMTD), and there we see that the profiles agree mostly very well. Thus we cannot attribute the differences to the broad coincidence criterion. We have added this argument to the paper.

Comment: *As it is, Figure 8 shows mostly a lot of disagreement with the balloon profiles but the reason is probably geophysical variation thus there is no point to these comparisons!*

Reply: This is a good point but as said before, we have made a similar comparisons for CFC-11 and CFC-12 (Eckert et al., AMT 2015) and methane (Laeng et al., AMT 2015). These comparisons are close to perfect. This refutes that the differences found can be attributed to geophysical variation. If they were, then the other gases should show similar (or even larger, due to larger vertical and/or horizontal gradients) discrepancies. This issue is now discussed in the revised version of the manuscript.

Comment: *And why calculate a 2005-2011 mean profile at all? This gas is increasing rapidly – at least 25% over this time period. The multi-year mean is meaningless, and should a balloon profile match it, that is meaningless too. Unless you can demonstrate that in spite of the broad coincidence criteria that it makes geophysical sense to compare with MIPAS, these comparisons could be eliminated.*

Reply: We agree that comparison to multi-annual means is not useful in this case and have eliminated them.

Comment: *Section 4.3, MkIV comparisons*
p. 14796, lines 15-28. Too much detail on the MkIV instrument use the Toon reference and eliminate most of this.

Reply: We consider some of the information as important. E.g. the fact that the same spectral band is used which is used for MIPAS may be important to interpret the comparison. However, we agree that there were too many details and we have shortened the paragraph.

Comment: *p. 14797. Since there are no MIPAS data for the dates of the balloon flights, these flights are not useful for validation! Again, comparing to a multi-year mean is not satisfying or meaningful in a quantitative way. This section can be eliminated.*

Reply: We have removed all lines and panels which are related to multi- annual mean data and restrict the discussion to the comparison of collocated measurements and measurements of the same month, latitude and year. For MkIV, there is still a meaningful pair of measurements available (23 Sep 2007). Further, the MkIV balloon observations were made in September after a summer of Easterly stratospheric winds. Since the flow is zonal during the summer (no wave activity) little zonal variation in composition is to be expected, which justifies comparison with a zonal monthly mean of the same year. We have used this approach for the 22 Sep 2007 data.

Comment: *Section 4.4 Summary of intercomparisons*
The summary states that, compared to the different data sets, the MIPAS bias is either low, high, or zero. This is not a result. It is exactly what was known before any comparisons were made.

Reply: We disagree. Now we know that there is no clear indication of a bias and that MIPAS does NOT stand out as particularly high or low. We did not know this before. We have reworded this statement for clarification in the paper.

Comment: *The balloon profiles really aren't adequate for this validation – that's fine, so don't use them.*

Reply: We disagree. The geolocation problem with the cryosampler data has been refuted by comparison of CFC-11, CFC-12 and CH4 data which rule out geophysical variability as explanation of the differences. Further, cryosampler data are the only stratospheric data set which does not rely on spectroscopic data and thus are the most independent reference data. For MKIV there is at least one coincidence available, and the comparison between MKIV and the MIPAS September mean of the same year has been justified above. Inclusion of multiple validation instruments has scientific value.

Comment: *The validation section would be greatly improved by expanding the ACE data comparisons as noted above. After a more thorough validation using the ACE data, I think you will be able to state much more definitively where the data sets agree and where there is bias (and how much).*

Reply: We agree that the ACE-FTS data are an excellent data source for comparison. Since, however, the current version of ACE-data used here have not yet been validated themselves, we consider it adequate to use also other data for MIPAS validation.

Comment: *Section 5. Climatology*
The data are presented in various ways (e.g., latitude v. time, altitude v. latitude, etc) but there is actually no climatology here.

Reply: We agree that 'climatology' was a misnomer for what we have presented. We have changed the title of the paper and the section header accordingly. We think that the chosen representations best support our conclusions. A adequacy of a classical climatology seems inadequate to us for a gas with a pronounced trend.

Comment: *In general, the analyses in the subsections are only descriptive (i.e., descriptions of what is already known) or speculative, and do not present any quantitative analyses.*

Reply: The content of Section 5 is both descriptive and explanatory. We think that also the descriptive part has scientific value, because it characterizes important aspects of the state of the atmosphere during the measurement period. We are reluctant to understand the term 'descriptive' in any pejorative sense here. A second issue in our discussion is the test of plausibility of the observed features as a part of data quality control. It is a necessary condition for considering the MIPAS data useful that they show the known features. In this sense we consider the cross-validation between MIPAS and available prior knowledge as useful and necessary. Thus, we do not see what is bad about the fact that MIPAS shows the known features. Closer analysis, however, reveals features which are prima facie unexpected. For these features we offer explanations, both to show that these features are not artefacts and because they hint at interesting processes. These explanations are morphological (Section 5.1) and quantitative (Section 5.2). Below, where the criticism is specified, we defend our explanations against the criticism of being speculative.

The fact that the mixing of monsoon air into the tropics is instantaneous and not restricted to the time when the monsoon breaks down is to our best knowledge a new finding.

The fact that the temporal development of HCFC-22 in the stratosphere cannot be explained by the trends measured at the surface and the known age of air induced time lag is also a new finding.

The fact that Asian HCFC-22 emission have become the major source of global upper tropospheric HCFC-22 has been plausible but our data are to our best knowledge the first empirical evidence in support of this hypothesis.

Comment: *No meteorological data are used in support of speculative statements about processes that might be indicated in the data. This section would be improved by including a climatology (i.e., mean distributions as a function of month/season, mean cycles, etc.) and by adding meteorological analyses to give support to the processes you describe.*

Reply: Climatologies are usually understood to be multi-annual monthly mean distributions along with their variabilities. For a gas with a pronounced trend we think that the presentation of climatologies in that sense is of limited use because this would imply averaging over an inhomogeneous data set. The representation in the paper has been chosen because it represents best the features we discuss.

Comment: *Figures 10 and 11 are introduced (p. 14798, line 3) but the next figure mentioned is Figure 15 (line 19). Figures must be mentioned in sequence.*

Reply: Instead of referring to Figure 15, we have now inserted an additional figure. This serves two purposes: First, we can leave (old) Fig 15 in the section on temporal development (where it belongs to). Second, the new figure dedicated to the monsoon issue better resolves what we want to show and makes our explanation more obvious.

Comment: *p. 14799, lines 7-9. While rapid uplift can explain why upper tropospheric values are nearly equal to surface values, they cannot explain mixing ratios that are higher than the surface. What would the source of the extra HCFC-22 be?*

Reply: Indeed we initially thought that there must something be wrong with our data. However, this is not true. The reason is roughly this. We look at 2D distributions. If there is a strong source of limited longitudinal extent somewhere, the excess vmr averages partly out when calculating the zonal mean. In a region, however, where the transport direction turns to zonal, the averaging goes along the transport direction, i.e. over a series of enhanced values, and the enhancement does no longer average out. Thus, in 2D we can see values which are larger than those in the source zonal band, although there is no other source further up in the atmosphere. In the 3D field we of course cannot have values larger than the source values at any place. Beyond this: We do not see East Asian surface values; surface values mentioned in this paragraph refer to clean air measurements where background concentrations are measured. So the apparent conflict is not even evident.

Comment: *Sections 5.1 and 5.2 have lots of qualitative discussion but there is no actual MIPAS data analysis that demonstrates any of the processes discussed; e.g., p. 14798, line 22: "The following scenario is suggested ...." Analyses, not suggestions, are required for publication. Why not bring in meteorological data to support your ideas?*

Reply: We admit that by using the term 'suggest' we undersell our analysis, thus we have reworded the related text. The fact that Asian countries' HCFC-22 emissions have increased is supported by a reference. The fact that there is uplift in the monsoon region, which would of course uplift also HCFC-22, is well established knowledge, and related references have been added. The literature referenced used a transport model driven by meteorological data, thus this request is now implicitly fulfilled. Indeed MIPAS HCFC-22 maps show a clear monsoon signal in the upper tropopause (newly included Figure, upper panel).

The flooding of the tropics and northern latitudes (the latter after the breakdown of the monsoon) with the excess HCFC-22 is clearly seen in the newly included Figure, lower panel. The fact that upper tropospheric transport and mixing in the tropics happens roughly in zonal direction is well established, and a reference has been added. Thus we see no gap in our chain of arguments and do not understand what is speculative about it. We admit that our chain of arguments in the original version of the paper appeared to have gaps without the references and the newly included figure.

Comment: *Also, Figure 15, which is used to show something about monsoon transport, crams 6 years of data on 5 surfaces into a very tiny space. The panels are illegible, the font impossible to read. It is impossible to see the details of a seasonal process such as monsoon transport with a tiny panel showing a 6-year time series.*

Reply: This is due to the landscape format of ACPD. In the final version, a full column will probably be available for these panels. The layout has not been chosen by us. Beside this, one can zoom into the figures in most PDF viewers. New figures have been provided.

Comment: *Other parts (e.g., first few paragraphs of 5.2) describe what is already very well understood about long-lived trace gas structures and seasonal cycles.*

Reply: We are not aware of any other published data set that actually contains information about the global altitude-resolved temporal development of HCFC-22 in the time period under investigation. The first part of Section 5.2. describes our data in the context of existing understanding of atmospheric processes. This strengthens the confidence both in our data and in the pre-existing assumptions. Given the novelty of our dataset in terms of global coverage along with temporal resolution, we consider it as appropriate to confront our a priori assumptions on atmospheric processes with the empirical data.

Comment: *Most air enters the stratosphere through the tropical tropopause, so interhemispheric (IH) differences in long-lived trace gases found in the troposphere are usually not found in the stratosphere. (CO2 has some but they disappear quickly with height.) If you speculate about IH differences in the stratosphere, you'll need a supporting analysis to demonstrate that tropospheric IH source differences are the cause.*

Reply: We do not agree. The interhemispheric differences of HCFC-22 are not explained by hemispherically different sources but by dynamical processes (e.g., polar seasonal cycle, semi-annual oscillation, QBO). For a source gas which has, contrary to CO_2, pronounced vertical and latitudinal mixing ratio gradients, any interhemispheric difference in stratospheric dynamics will cause differences in the HCFC-22 distributions. Over the poles, the mixing ratio changes are mainly driven by subsidence of HCFC-22-poor air from above. Later, this air is mixed with HCFC-22-rich air from mid-latitudes. It is well known that interhemispheric differences with respect to these processes do exist.

Comment: *p. 14800, line 14. 'Interestingly, the breakup of the vortex seems to take place at all altitudes at almost the same time for the northern polar region.' The Arctic vortex final warming ('breakup') occurs in March or April and shows considerable interannual variability in how it breaks up (e.g., wave 1 or wave 2 warming). Its variability is much greater than the Antarctic vortex, so your statement doesn't make sense. The sharp discontinuities found each year in your figure appear too early in winter to be the breakup they are probably midwinter sudden warmings, not the final warming (breakup) that occurs in March or April. Please check your figures and interpretation.*

Reply: The 'same time' is not meant in an interannual sense but intends to say 'at the same time at all altitudes'. The wording in the manuscript has been changed for clarity. The discussion of the warming events has been revised. Indeed we have missed an important point: The sharp increase of HCFC-22 in the zonal means does not require a vortex breakup but can, in the 2D representation, also be caused by deformation or displacement of the polar vortex, i.e. by everything which leads to averaging over vortex and non-vortex air when the zonal mean is calculated. The early HCFC-22 increases in the northern polar winters just reflect this wave activity. Thanks a lot for attracting our attention to this issue.

Comment: *The lowermost stratosphere is generally below 16, not 20 km. It is below 380 K (below Hoskins' 'overworld').*

Reply: The wording has been changed.

Comment: *p. 14801, line 15. 'the seasonal cycle in the SH ... is not globally compensated but only weakened by the cycle of the NH': What does globally compensated mean?*

Reply: If NH and SH signals were exactly the same but out of phase by half a year, then they would cancel (compensate) on a global scale to zero annual cycle. The related text has been reworded for clarity.

Comment: *The results of the growth rate analysis in Section 5.3 (p. 14803) might be better displayed as a table. Too many different units are used to discuss growth rate, making it harder to compare results with previous studies cited. The Figures (16,17) use ppt/yr and%/decade. Studies cited on page 14803 use*

%/yr. Pick a unit and stick with it.

Reply: A table has been added; units have been unified in all cases when we knew the reference data used for the calculation of the percentage growth rates.

Comment: *Section 5.4 Comparisons with surface measurements The NOAA/ GMD and AGAGE measurements are both precise and accurate: why compare in two separate sections? Please combine them into a single 'comparisons with surface'*

Reply In the original submission these data sets are already treated in one section (5.4). In the revised version we have combined the NOAA and AGAGE data sets in the figures, and have sorted the figures according to latitude bands, as suggested later in this review.

Comment: *section. p. 14804. 'the troposphere can be considered well-mixed'. Yes, sort of. As you can see in your own figures there are interhemispheric gradients of 20 ppt (10%) at the surface.*

Reply: The wording has been changed towards 'free troposphere' and 'within each hemisphere', and 'vertically' has been inserted, since this is what we intended to say.

Comment: *Why is the first paragraph of Section 5.4 talking about the dryness of stratospheric measurements when this section is about comparisons (of dry mole fraction) ground-based measurements with MIPAS tropospheric measurements? You haven't explained whether the MIPAS tropospheric measurements are wet and whether this will impact the comparisons with ground-based data. If there is an impact, what size is it? Big enough to cause a bias, or is it in the noise compared to other differences?*

Reply: Where MIPAS measures, the air is so dry that the difference between dry air mole fraction and actual mole fraction is not an issue. Any humidity which would cause any discernable difference between dry air and wet air mixing ratios requires so much water vapour that the water vapour signal would dominate the measurements and no HCFC-22 retrieval would be possible at all. Thus, in the context of MIPAS measurements, the distinction between dry air and wet air mixing ratios is meaningless because the 'wet' air is still dry enough. Text w.r.t. this has been added.

Comment: *p. 14805. If MIPAS measurements and ground measurements are the same to within their overlapping uncertainties, then this means they AGREE. Say so.*

Reply: Isn't that what we say on top of page 14805?

Comment: *Figures 18 and 19 have too much data. It is hard to distinguish MIPAS behavior at different latitudes. I suggest organizing the comparisons into panels showing different latitude ranges (e.g., NH extratropics, tropics, and SH extratropics) and include the relevant GMD and AGAGE stations on each panel.*

Reply: Agreed and done.

Comment: *p. 14805, lines 21-24. Again, this is more speculative discussion that is not supported by meteorological analyses or other trace gas measurements. This section could be very interesting if you examined the HCFC-22 data as a function of altitude in the lower-most stratosphere (e.g., 12-16 km) in each hemisphere to identify seasonal transport processes in the lowermost stratosphere. Interhemispheric differences between the results may reveal process important to the cycles in each hemisphere.*

Reply: We do not see the gap in out chain of arguments. The fact that NH and SH air is mixed in the tropical convergence zone or at least close to the stratospheric entry point is established knowledge, and it is even used in the review to refute our explanation of stratospheric interhemispheric differences. Also the fact that the outflow of the Hadley cell takes place in the upper troposphere is well established knowledge. From this, we can deduce that the interhemispheric contrast must be smaller there than near surface. We cannot recognize the speculative component in the deductive-nomological explanation.

We admit, however, that there was an error in our text. Of course we do not mean the tropical pipe but the inter-tropical convergence zone.

We agree that there is room for a more thorough discussion of this issue, but this involves further species. We think think that a more thorough discussion of this issue shall not be made in a paper which focuses on a single species, because such an analysis is much better made using the complete suite of tracers provided by MIPAS.

Given the altitude resolution of MIPAS (3 km or worse), we consider it questionable if a lot can be learned from the examination of HCFC-22 as a function of altitude between 12 and 16 km.

Comment: *p. 14806. I see that you recognize that chemistry, emissions, and transport are all important to understanding HCFC-22 behavior. But this means that it cant be understood as simply as discussed here. An atmospheric model is required to adequately interpret the behavior.*

Reply: We would agree if we proposed new hypotheses here. What we do, however, is only to put our results in the context of existing explanations found in the literature. I think it is fair to state that our results support these findings, and for this we do not need dedicated model calculations of our own.

Comment: *Section 5.5. Stratospheric trends*
This paper examines a 7-year data set. Seven years is less than 3 QBO cycles. Each QBO cycle is different in length and its seasonal timing. It is not completely accounted for in the regression analysis by considering terms at 2 pressure levels.

Reply: By using an empirical QBO proxy we do not need to assume that the QBO is periodical. Beyond this, our regression analysis tool takes correlated residuals into account. Thus any systematic residual is accounted for in the respective error analysis.

Comment: *Because the regression analysis cannot adequately remove the QBO effect, the residuals (Figs. 17 and 20) have a QBO signature. To search for a trend (that indicates stratospheric circulation change) in a 7-year data set is by definition unreasonable.*

Reply: We disagree. If there was a significant residual QBO signal, this should be visible in Fig. 16. There, however, no systematic QBO signal is discernable, and the residuals are small (2% of the amplitude of the largest periodic). This is no wonder because the QBO is part of our regression model. Of course a short time series is no adequate data base for a climatological trend analysis. But that is not what we do. We analyze the linear component of the temporal development (which we call 'trend') within a time window and use it to analyze the atmosphere within the same time window. We do not extrapolate any trend nor do we use it in any climatological sense. Thus, it is not relevant how long the time series is. All uncertainties are included in the error estimation of the trends.

Comment: *There are end effects (2 years out of the 7 are endpoints!), the QBO cannot be adequately accounted for, and in the case of HCFC-22, there is an enormous annual growth (compared to a possible circulation change) in the species measured.*

Reply: This argument suggests that there are only seven data points. This is not true: We use monthly data, i.e. appr. 12 times 7 = 84 data points. That means that we have 82 data points that are not end points. QBO and seasonal cycle are included in the regression model. Thus, they cannot alias into the trend. Ambiguities between these coefficients due to non-orthogonality are considered in the error estimation and subsequent significance analysis.

With respect to the enormous annual growth we are not sure if we understand the argument of the reviewer correctly. Is the fact that differences of two large numbers are usually associated with a large error of the difference considered to be the weak point? If so, we defend our approach as follows: The uncertainties of our trends are calculated along with the trend estimation itself where autocorrelations are considered. The uncertainties are about 1-2% per decade. The growth rates from surface measurements are considered very accurate, and besides this, related errors are considered as irrelevant when differences between unexplained stratospheric MIPAS trends between different latitude/altitude bins are analyzed. The unexplained part of the MIPAS trends ranges between -40%/dec to +40%/dec, which is far beyond the uncertainty. From this we conclude, that our estimates are still significant.

Comment: *Section 5.5 should be eliminated.*

Reply: Since related criticism has been refuted, we have kept this section.

Anonymous Referee #2:

Comment: *The paper by Chirkov et al. provides important information about HCFC-22 ($CHClF_2$) data in the whole stratosphere and upper troposphere, as derived from MIPAS (Michelson Interferometer for Passive Atmospheric Soundings) global observations performed in the "reduced resolution mode" over a little more than 7 years, starting in January 2005. Several aspects are covered, from a brief description of the retrieval to the determination of the global distribution of HCFC-22 and the changes in its concentration with time and altitude over the available years.*

Potentially, this is an important contribution for a "Montreal Protocol species" which is poorly sampled in the upper atmosphere, with global measurements only available from ACE-FTS since the loss of the Envisat satellite three years ago. The paper fits well with the scope of Atmospheric Chemistry and Physics but includes several annoying imperfections which should have been corrected by the authors or spotted by the editor before submission or online publication. I would therefore recommend publication after some significant reorganization and rewording, also considering the suggested changes outlined or detailed below.

General comments
The current title includes the words "climatologies" and "trends". This clearly corresponds to two overstatements in a row for a data set covering 7 years or so. To avoid any misunderstanding, I recommend changing the title to something like "HCFC-22 measurements with MIPAS: retrieval, validation, global distribution and its evolution over 2005-2012".

Reply: This is an excellent idea. The title has been changed in order not to raise false expectations. Beyond this, the term 'climatology' is no more used in the paper but more specific terms are used. The term 'trend' makes more problems, because any attempt to avoid this term leads to cumbersome language. We have removed the term from the title and most section headers (except one). In the text a statement has been included that the term trend as used in this paper is not meant in any climatological sense but describes only the linear component of the temporal variation within the time window under analysis.

Comment: *This data set is very important for the scientific community. Beside the discussions, the current presentation is essentially restricted to a suite of (sometimes small!)...*

Reply: As said above,in the reply to reviewer #1, small figures occasionally are caused by the landscape format of the Discussion version of the paper. In the final publication a full column will be available for these figures.

Comment: *...color plots which will be of limited use to the interested reader. I would therefore strongly suggest to include the most important information in an electronic supplement, as done e.g. in Kellmann et al., ACP, 12, 2012. This supplement should at least include the underlying data used to build the color plots (starting Figure 10) and the time series of Fig. 16, allowing direct numerical comparison with model outputs, computation of "trends"*

Reply: Trend data is now made available as a supplement. The original HCFC-22 data are available via our data server. A link is now provided in the paper.

Comment: *Figure 12 presents interesting results showing similarities with material published recently for other stratospheric tracers, i.e. in Nedoluha et al. (doi:10.5194/acp-15- 6817-2015, see Fig. 10) and Mahieu et al. (doi:10.1038/nature13857, see Fig. 4). Wouldn't this be helpful when addressing the "HCFC-22 unexplained relative trend"? A brief discussion putting these findings into perspective is welcome in section 5.5.*

Reply: This has been included in the discussion

Comment: *Please, also consider the following suggestions and corrections: Page 14785-L7: the modeled spectrum is fitted to the observation, not the opposite! Change to "...fitting of the modelled spectra to measured limb spectral radiances".*

Reply: Agreed and corrected.

Comment: *Page 14786-L7: replace CHF_2Cl by $CHClF_2$ to conform to the IUPAC nomenclature of organic chemistry (i.e. here alphabetical ordering of the substituents).*

Reply: Agreed and done. But what about OH? :-)

Comment: *Page 14786-L15: I believe that the IPCC assessment (the so-called "AR-5") should also be cited here.*

Reply: Agreed and done.

Comment: *Page 14786-L18: the correct word for 2007 is "Adjustment", not "Amendment". So update to "The 2007 Adjustment to the Protocol..."*

Reply: Agreed and corrected.

Comment: *Page 14787-L20: ground-based might be misleading here, I suggest "from surface long-term data records"*

Reply: Agreed and done.

Comment: *Sections 3 and 3.1: even if the information is available from the references you are citing, you need to mention here the actual line or cross-section parameters adopted in the MIPAS retrieval scheme for the target and interfering species!*

Reply: Agreed and done.

Comment: *Page 14790-L2: suggest changing to "....the sole contribution of HCFC-22 is shown in red."*

Reply: Agreed and done

Comment: *Page 14790-L21: you are retrieving HCFC-22 from 7 years of observations, and the error budget provided in Table 1 corresponds to a single observation. How could this be? At the very least, we need to know if these numbers/figures are representative/typical, or correspond to a "best-case". E.g., do you see a significant scatter among the individual/per orbit error evaluations? This is also important in view of the comparisons with other instruments presented in section 4.*

Reply: This error estimate is considered to be roughly representative for the entire data set but rather on the conservative side, due to the typically low lower stratospheric temperatures at tropical latitudes, which go along with a lower signal. A note on this has been included in the paper.

Comment: *Section 4.1: here also, you have to mention the origin of the line parameters used for the ACE retrievals. Different line parameters could lead to systematic biases. A proper validation exercise requires this kind of information.*

Reply: Good point! This information has been included for all spectrometers used in this study.

Comment: *Section 4.3: Same remark as for sections 3.1 and 4.1 (spectroscopy). Also, the MIPAS, ACE and MkIV retrievals use dissimilar windows. What about the possible impact of these choices on the (validation) results? This should be quoted.*

Reply: The windows are somewhat different, but this is basically to compensate the effects due to different spectral resolutions and different retrieval settings. Even with the same microwindows for all spectroscopic instruments the results are expected to vary, perhaps even more compared to the case when each instrument uses microwindows custom-tailored to the specific needs of the instrument and retrieval scheme.

Comment: *Page 14796-L25: change to "on a 1 km".*

Reply: Agreed and done

Comment: *Page 14796-L27: change to "Fort Sumner, NM".*

Reply: Agreed and done

Comment: *Section 4.4: The statistics of the comparisons (probably a word more appropriate than "validation" in the context of this paper) are extremely different. Only a handful MkIV or cryosampler flights are presented (btw involving MIPAS means...) when more than 8000 collocated measurements with ACE have been used! It is unclear to me whether this is properly accounted for in the concluding remarks of section 4.4.*

Reply: There are two aspects. With respect to noise, the ACE intercomparison is of course much more significant. However, there is not only noise but there are also systematic issues. These do not cancel out. This is why we also consider reference instruments which provide only a small number of correlated measurements. A note on this has been included in the manuscript.

Comment: *Section 5.1: This section starts with a brief description of Figures 10 and 11. Then suddenly, on line 19, you discuss about the results of Fig. 14* *(compare Fig. 14, but with what?) and of Figure 15 on next line. These figures have not been described nor introduced in the text and they are mentioned before Fig. 12 and 13. This needs to be seriously revamped, eventually involving a new ordering of the figures and/or sections.*

Reply: We have included an additional figure. This serves two purposes. First, the figures are called in proper sequence. Second, the new figure better supports the discussion in Section 5.1

Comment: *Page 14799-L16: "compare Fig. 15, panel 2", what do you mean here? Do we need to compare panel 2 with the other ones? With another figure of the present paper, or of another paper? Please specify.*

Reply: It was meant to check panel 2 of Fig. 15. However, a new figure has been included and the wording has been changed.

Comment: *Section 5.4: an alternative title might be "Comparisons between tropospheric and surface growth rates".*

Reply: This sounds indeed much better, thanks. We have modified the title to "upper tropospheric"

Comment: *Page 14804-L6: change to "do not reach the ground".*

Reply: Agreed and changed.

Comment: *Page 14816-L11: change CHCLF2 to CHClF2.*

Reply: Agreed and corrected.

Comment: *Figure 8: the two different green curves are hard to distinguish once printed.*

Reply: The multi-annual mean has been removed. Thus, only one green curve remains.

Comment:*Figure 9: are the thin curves on the lower panel identified in the legend?*

Reply: The error bars have now been identified in the figure caption. Or is the problem with the grey curves? These are all MIPAS profiles within the collocation radius around the MkIV geolocation. The legend says "MIPAS collocated".

Comment: *Figure 15: the panels are really small, it would probably be preferable to arrange them as in Figure 13.*

Reply: In the ACP version the figures are printed in portrait instead of landscape format. Thus the figures will be larger.

References:

Brown, A.T., et al., Trends in atmospheric halogen containing gases since 2004, J. Quant. Spectrosc. Rad. Trans. 112, 2552-2566, 2011.

Brown, A.T., et al., Stratospheric lifetimes of CFC-12, CCl4, CH4, CH3Cl and N2O from measurements made by the Atmospheric Chemistry Experiment-Fourier Transform Spectrometer (ACE-FTS), Atmos. Chem. Phys. 13, 6921-6950, 2013.

Brown, A.T., et al., Global stratospheric fluorine inventories for 2004-2009 from Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS) measurements, Atmos. Chem. Phys. 14, 267-282, 2014.

Eckert, E., et al., MIPAS IMK/IAA CFC-11 (CCl3F) and CFC-12 (CCl2F2) measurements: accuracy, precision and long-term stability, Atmos. Meas. Tech. Discuss., 8, 7573-7662, doi:10.5194/amtd-8-7573-2015, 2015.

Kellmann, S., et al., Global CFC-11 (CCl3F) and CFC-12 (CCl2F2) measurements with the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS): retrieval, climatologies and trends, Atmos. Chem. Phys., 12, 11857-11875, doi:10.5194/acp-12-11857-2012, 2012.

Laeng, A., et al., Validation of MIPAS IMK/IAA methane profiles, Atmos. Meas. Tech. Discuss., 8, 5565-5590, doi:10.5194/amtd-8-5565-2015, 2015.

Nassar, R., et al., A global inventory of stratospheric flourine in 2004 based on Atmospheric Chemistry Experiment Fourier transform spectrometer (ACE-FTS) measurements, J. Geophys. Res. 111, D22313, 2006a.
Nassar, R., et al., A global inventory of stratospheric chlorine in 2004, J. Geophys. Res. 111, D22312, 2006b.

Rinsland, C.P., et al., Atmospheric Chemistry Experiment (ACE) Arctic stratospheric measurements of NOx during February and March 2004: Impact of intense solar flares, Geophys. Res. Lett. 32, L16S05, 2005.

von Clarmann, T., et al., Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), J. Geophys. Res., 108, 4736, doi:10.1029/2003JD003602, D23, 2003.

von Clarmann, T., et al., Retrieval of temperature, H2O, O3, HNO3, CH4, N2O, ClONO2 and ClO from MIPAS reduced resolution nominal mode limb emission measurements, Atmos. Meas. Tech., 2, 159-175, doi:10.5194/amt-2-159-2009, 2009.