Responses to referee#2

We wish to thank the referee for his review. Our responses are given below where text in black corresponds to the referee comments, text in blue corresponds our responses and text in green corresponds to the text of the revised manuscript. All page and line numbers provide here refer to the manuscript available on the AMTD page, not to the upcoming version of the manuscript.

In addition to the change made in the manuscript to take into account your comments or the comments of the other referee, several other changes have been made and are listed here.

P4L90: "MIPAS" replaced by "MIPAS ESA"

P12L368-L373: "The discontinuities .... may occur" replaced by

> "These discontinuities are due to the combination of two effects. The first one is due to the decontamination procedure of the instrument (i.e. a warming-up of the instrument to remove the ice) which is operated once or twice per year. Sometimes after the decontamination an abrupt change, as high as 2%, is observed in the radiometric gain of band B where CH_4 and N_2O are retrieved. The second effect is due to the calibration of the L1 data which is done once a week. Since the change in the gain occurs in the timeframe of 1-2 days, and since the calibration is not performed at the corresponding times, a discontinuity in CH_4 and N_2O time series is introduced. This issue should be resolved in the future version 8 of MIPAS by doing daily calibration of L1 data."

P12L382 and P13L418: "weekly calibration of L1 data" replaced by "abrupt change in the radiometric gain"

General Comments

In this paper, Errera et al. describe assimilation experiments using MIPAS upper-tropospheric/stratospheric CH4 and N2O profiles from the ESA processing version 6 and, partly, also version 7. These tests indicate deficiencies in the retrieval results in the equatorial lower stratosphere/upper troposphere and problems due to the calibration during certain periods. The aim of the study is to demonstrate the possibilities of data assimilation for the diagnosis of space-borne atmospheric observations. To make this point clearer, it would be good if the authors could summarize, e.g. in the conclusions, which issues could only (or with more confidence) be detected due to the application of data assimilation and which would be obvious also without this technique. After addressing the items listed below, I support publication of the manuscript in AMT.

The conclusion has been updated in that sense and also to take into account a comment of referee#1. The 2nd § of the conclusions (P13L417-L422) has been rewritten as:

> "Nevertheless, this study also diagnoses two issues in MIPAS CH_4 and N_2O profiles. First, time series of MIPAS profiles show unexpected discontinuities which are due to the abrupt change in the radiometric gain of the instrument. A daily calibration might resolve this issue. While identified in this paper, this issue could have been found by data analysis methods other than data assimilation. Second, the correlations between BASCOE analyses and independent observations from MLS and ACE-FTS are

poor in the tropical lower stratosphere. This is due to outlier profiles which are not flagged out in the presence of clouds. This second issue was not identified in previous validation studies of MIPAS. One possible reason is that method to compare satellite observations usually depend on a coincidence criteria that limits the size of the sample. This is not the case any more if one compares data assimilated fields from one satellite instrument with observations from another satellite instrument as done in this paper. These two issues are also present in MIPAS version 7 and will be addressed in the future version 8."

P13L432 sentence "Finally, this study ..." is removed.

Specific comments

P1L5: 'can be noisy'. The term 'noisy' may imply that the effects are due to measurement noise; perhaps 'unstable' would be more suited(?)

Indeed, "noisy" might not be appropriate. We will replace it with "vertically oscillating" because it describes the shape of the profiles without indicating a cause for it.

P1L5: "noisy" replaced by "oscillating".

P1L6: 'B matrix'. This is 'slang' for insiders. Please use an explanatory term.

The term "B matrix" has been removed from the abstract but used in the body of the paper since well accepted by the data assimilation community (e.g. Bannister et al, 2008).

P1L11: 'the calibration method'. Is it really the method or the availability of calibration data/files?

The discontinuities are due to an abrupt change in the gain and to the choice of performing the calibration weekly, even though daily calibration is possible. In V8 daily calibration data will be used and the impact of the abrupt change in the gain will be reduced. For V7, it is not a problem of availability of data/files, but a problem of performing the calibration once per week. The text has been modified to better explain this. P1L11 has been changed accordingly:

"First, time-series of the observations show unexpected discontinuities, due to an abrupt change in the gain of MIPAS band B, generally occurring after the instrument decontamination. Since the calibration is performed weekly, the abrupt change in the gain affects the measurements until the subsequent calibration is performed."

P4L92: '(NOM) with altitude soundings between 7-72 km'. I think it is important to note here the latitude-dependence of the lowest tangent altitude in the OR phase.

In the revised paper we have pointed out that in the nominal mode of the OR phase the measurement grid depends on the latitude following a floating-altitude law. P4L93 added before "Most of the MIPAS profiles ...":

"In OR measurements, NOM mode is characterized by a floating-altitude measurement grid. This means that the limb sounding grid is shifted rigidly with the lowest measured altitude which depends on the latitude. The floating-altitude sampling grid is meant to follow roughly the tropopause height along the orbit with the requirement to collect at least one spectrum within the troposphere while avoiding too many spectra affected by clouds."

P4L105-120: Please give the exact pages, where to find this in Rodgers, 2000. How has the interpolation of A and y0 been performed? Please give the formulas. Which vertical-grid has been used (altitude, pressure, theta)? Have dedicated AKs for each MIPAS profile been applied?

The equation is found more readily in Rodgers and Connor, 2003, JGR, now cited in place of Rodgers (2000).

The interpolation of A and y0 is linear along the logarithm of the pressure. $\tilde{A}$ is normalized such that sum(A_i)=sum($\tilde{A}\_i$), where *i* correspond to a row of *A* or $\tilde{A}$. Is there really a need to give a formula for that? MIPAS provides AK for each profiles which have been used as stated L105. P4L119 now has a new sentence:

> "The interpolation is done along the logarithm of the pressure and $\tilde{A}$ is normalised in order to have $\sum_j A_{ij} = \sum_j \tilde{A}_{ij}$ where *i* and *j* denote, respectively, the rows and lines of A or $\tilde{A}$."

P5L129: 'This is due to the use of an oscillating MIPAS profile'. Is it really only due to the oscillating MIPAS profile or also due to the AK? Could you please explain why it is 'allowed' to apply the Rodgers formulation using an already oscillating yk – should the formulation not also be valid when a non-oscillating profile is used as yk (I understood that this is not the case)? If it is only due to the profile and the oscillations are not in the AK, does this mean that the reason for the oscillations are due to some kind of systematic error (comparing the error bars in Fig. 1b with the strength of the oscillations it seems that those are not only due to random errors)? Is there any speculation on the reason for the oscillations? Will those be less strong in MIPAS version 8?

We can see in Fig. 1 that the AKs are positive peaked functions with different widths at different altitudes. Therefore, we do not see how the AKs could introduce oscillations in the result of Eq. (2). The Rodgers formulation should be valid in both cases of oscillating or non-oscillating profiles yk. We agree with the reviewer that since the random errors are smaller than the oscillations, some systematic errors should induce these oscillations. At the moment we have not identified the cause of these oscillations but further investigations are ongoing.

P7L221: 'the variance of R is given by the ML2PP retrieval'. Does this account only for random noise or also for 'random'-like or systematic other error components, like calibration etc?

The R given by the ML2PP retrieval accounts only for random noise. We have specified this in the revised paper. P7L221 is replaced by:

> "In this paper, the variance of R (which accounts only for the random noise errors) is given by the ML2PP retrieval …."

P8L253: 'Figure 3 shows the zonal mean analysis of CH4 from six BASCOE experiments'. It would be interesting to plot here also zonal means of the 'pure' MIPAS dataset. Further, has the latitude dependent lower boundary of MIPAS been considered? I.e. are you sure that no data are plotted which are below the lowest tangent altitude of the limb-sounder or is this covered by use of location-specific AK's?

Zonal mean of MIPAS CH4 will be added in Figure 3 (see below) with white spots to indicate missing values. MIPAS zonal mean are from MIPAS data that are binned on a 5 degree

latitude grid and the a pressure grid with 6 pressure levels per decade of pressure. The latitude dependence of MIPAS w.r.t. the lower boundary is considered since the altitude of MIPAS tangent point is identified by the pressure profile, which is itself measured by MIPAS.

Here is the new Fig. 3 with the updated legend:

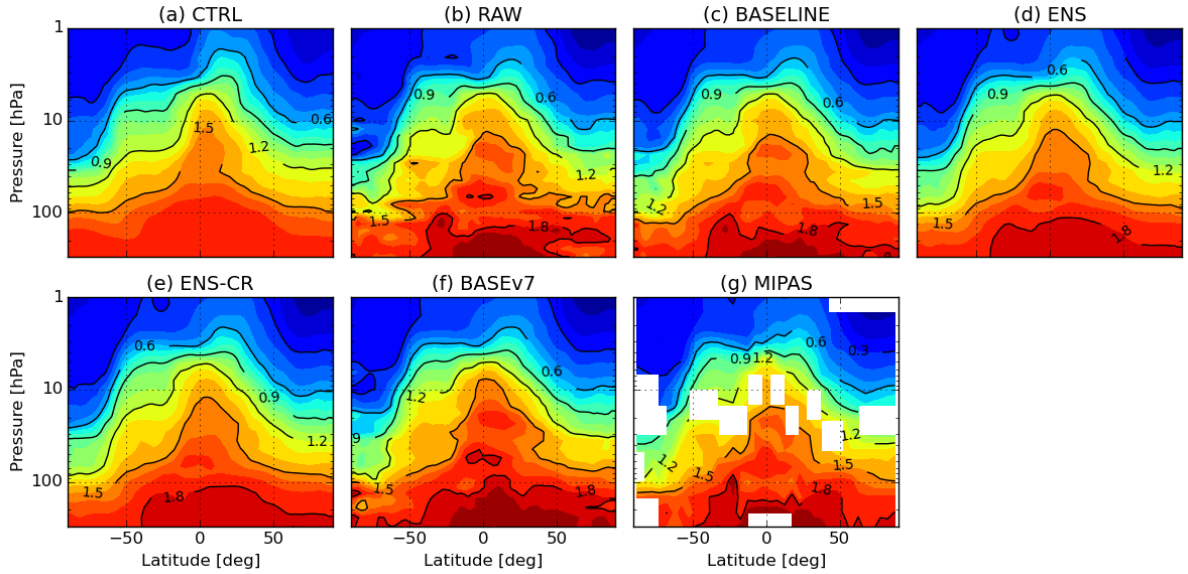

Figure 3: Zonal mean of CH4 (ppmv) from six BASCOE experiments and from MIPAS on September 15, 2008 at 12 UT (see text for details). Here MIPAS data are binned on a 5 degree latitude grid and a pressure grid with 6 pressure levels per decade of pressure.

P8L253, added after "is representative of other dates.":

"The MIPAS CH_4 zonal mean is also shown in the figure in order to allow comparison with the assimilated dataset."

P9L293: 'This suggests that the observational error covariance matrices provided by the MIPAS ML2PP retrieval are not optimal for data assimilation.' Could you show such a covariance matrix? Are there any ideas if there are general problems with the error covariances of MIPAS?

See answer to referee 1 regarding this issue.

P12L384: 'in the tropics. In that region, REAN analyses are relatively noisy and the N2O seasonal variations observed by MLS are not reproduced by the reanalysis of MIPAS N2O and CH4.' Do you mean at the lowest level? Any idea why this is not observed by MIPAS? Is it present in the MIPAS raw data? Could it be due to cloud contamination? Is this feature present in the raw simulations?

We are referring to the lowest level. Seasonal variations observed by MLS are hardly seen in the MIPAS raw data more likely due to cloud contamination in the tropics (as mentioned in the § starting at P12L386). It is not clear if this feature is present in the RAW simulation since it only covers the period April-November 2008. The following sentence will be added at P12L385:

"Note that seasonal variations observed by MLS are hardly seen in the MIPAS raw data."

P13L415: 'are in good agreement'. This expression is also used elsewhere in the paper. It would be better to make 'good' more quantitative by e.g. providing a specific context ('good' compared to what?).

The sentence is changed by: "... measured by MLS and ACE-FTS, usually within ±10%".

P13L418: 'show unexpected discontinuities which are to be due to the weekly calibration of L1 data'. It would be very helpful for users of MIPAS data to provide a table with the exact dates affected by the calibration issues.

This table is already reported in Table 5 in De Laurentis and Raspollini (2016) available here: https://earth.esa.int/documents/700255/2635669/RMF_0141+MIP_NL__2P_issue1.pdf/59beb833-5ad4-4301-8422-f41001da36d4 . A reference to this document is provided in the revised version of the paper.

The following sentence is added at the end ofP12L384:

> "Days for which L1 data are not properly calibrated are reported in De Laurentis and Raspollini (2016, see Table 5)."

P21Fig5: It would be illustrative to show the ACE-FTS regression curve from (a) in all subplots (b-f).

Done, see the updated figure and its caption.

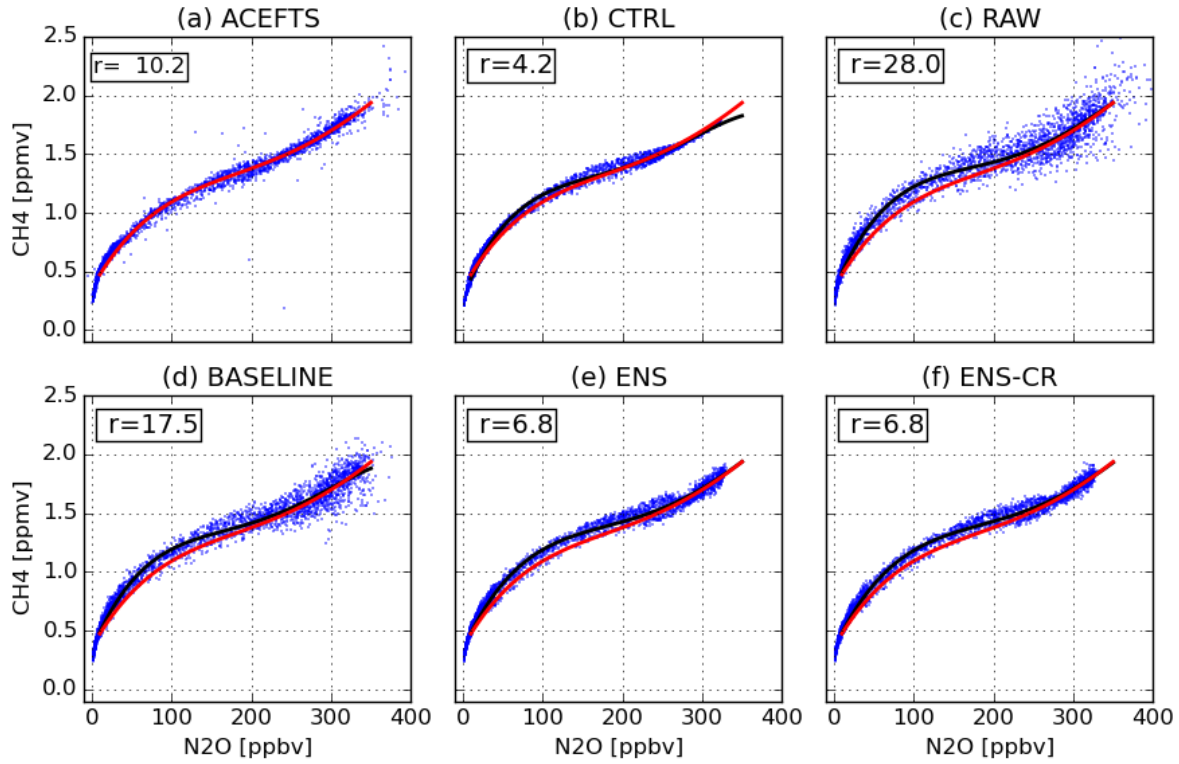

Figure 5. N_2O-CH_4 scatter-plot between 30°S-30°N as observed by ACE-FTS (a) and for five BASCOE experiments interpolated to the ACE-FTS observation space (b-f). The correlations are shown for the period April-October 2008. The result of a four-degree polynomial fit is also shown (red line for ACE-FTS, black lines for BASCOE experiments) with the root mean square of the residual between the fitted curve and the observations printed in the upper left of each plot. The correlation curve of ACE-

FTS (red curve) is reported in the correlation plots of the BASCOE experiments (b-f).

Technical corrections

P6L171: ‘recommandation’ -> ‘recommendation’

Corrected

P10L326: ‘ (Fig. 5)’: but BASEv7 is not shown there(?)

P10L326 “correlation is better with BASEv7 (Fig. 5)” replaced by “correlation is better with BASEv7”

P11L338: 2x ‘agreement’: please improve text

Corrected

P21, Figure 5a: please use the same grid lines as in other sub-plots P21, Figure 5 caption: ‘expriments’ -> ‘experiments’

Corrected, see above.