Responses to referee#1

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

6   In addition to the change made in the manuscript to comply with your comments or the
7   comments of the other referee, several other changes have been made and are listed here.

8   P4L90: "MIPAS" replaced by "MIPAS ESA"

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

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

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

22

23   1. General comments

24   The authors are having a close look at the data from the Michelson Interferometer for
25   Passive Atmospheric Sounding (MIPAS) as retrieved by the ESA processor (versions 6 and 7).
26   They use data assimilation to (i) fill spatial and temporal gaps in the data, (ii) smooth the
27   data using both averaging kernel information and a chemical and transport model, and (iii)
28   compare the MIPAS data with other data sets (ACE-FTS for CH_4 and N_2O and MLS for N_2O).
29   The authors are using different configurations of their data assimilation system BASCOE
30   (Belgian Assimilation System for Chemical ObsErvations). This helps them to better
31   characterise the MIPAS data as the effect of the assimilation could be assessed through the
32   ensemble of assimilation configurations. In particular, they provide useful information for
33   the possible users of this MIPAS product, e.g. their recommendation to use the ESA MIPAS
34   data (versions 6 and 7) with caution in the lower stratosphere over the tropics.

35   I think this study provides a good insight of the ESA MIPAS data (versions 6 and 7) and the
36   ability of BASCOE to assimilate them and provide a better product than the data alone. For
37   that reason, I would advise to publish it in Atmospheric Measurement Techniques with
38   minor changes listed hereafter.

39   Nevertheless, my main comment is that there is a mismatch between the introduction and
40   the options the authors took to carry out this study. My reading of the introduction is that
41   chemical transport models like BASCOE are the only alternative to study the stratospheric
42   chemistry as resolving "state-of-the art chemical equation systems for the stratosphere is

43 much too expensive for NWP models" (see comments below about this statement).
44 Curiously this study uses a version of BASCOE without any stratospheric chemistry! The
45 impact of this choice is not even discussed. I think part of the introduction should be
46 rewritten and the impact of not using stratospheric chemistry for CH_4 and N_2O in BASCOE
47 should be more discussed.

48 The chemistry of CH_4 and N_2O must be taken into account for long model simulations (>1
49 year). The assimilation of CH_4 and N_2O with the chemistry turned off is possible because the
50 chemical processes that affect these species occurs on time scales which are much longer
51 than the length of the assimilation window (24h), as stated in P6L183-L187. The manuscript
52 has been clarified. See also our responses to your specific comments below.

53 We inserted at the end of P3L73:

54 "All the assimilation experiments performed in this study have been done with the
55 chemical scheme of BASCOE turned off. This is possible since the chemical processes
56 that affect CH_4 and N_2O occur on time scales which are much longer than the 24h
57 assimilation window used in this study."

58 My second concern is about the assimilation experiment referred to as ENS-CR. In this
59 experiment the correlation of the observation errors are accounted for. It is well-know that
60 for any assimilation system assuming zero correlation in the observation error when the
61 correlation exists, one has to "tuned" the observation error variance. As a consequence, re-
62 introducing the observation error correlation reduce the weight of the observations and the
63 observation error variance should be "re-tuned". This was not done in this study where the
64 observation error variance is the same for the assimilation experiment with and without
65 observation error correlation. Figure 6 is a good illustration that the observation error
66 variance should be changed for the ENS-CR experiment. I would recommend to have
67 another ENS-CR experiment with an adjusted observation error variance to make more
68 sense of this experiment.

69 In all experiments, the observation errors were not tuned but taken directly from the ML2PP
70 retrieval with the caveat that the relative error is not allowed to be lower than 5%, so as to
71 account for the error of representativeness. The reviewer has a valid point that the effective
72 observation error variance is actually increased when vertical correlation of errors is present
73 as in ENS-CR experiment. Consequently and as seen with the $J(x_a)/p$ diagnostic the
74 observation error variance is perhaps too small, since $J(x_a)/p$ is well above the value ½. To
75 produce a recalibrated observation error variance poses the question on how to do this. One
76 could introduce a scaling factor for all variances at all levels, or adding an error variance
77 throughout, which effectively would result in increasing the lower limit of 5% for the
78 representativeness error. Observation error could also be tuned separately for each vertical
79 level using for example a Desroziers' method, but this would have introduced a whole new
80 set of calibration procedures. Additionally one could also argue that the observation weight
81 is incorrect in all experiments which use an uncalibrated **B**. Thus we do not pretend that we
82 compare experiments with using correct observation weights. Rather the paper is about the
83 sensitivity of [the results? the analyses?] to different configurations, and show how these
84 different configuration elements could be brought up together to provide a better analysis.

85 The text describing experiments ENS-CR has been rephrased in that sense, see in the
86 responses to your specific comment below.

87

88 2. Specific comments

89 P1L5: "The CH_4 and N_2O profiles can be noisy". I am still not convinced by this statement
90 after reading the whole text as there is no figure to really prove this or we cannot get easily
91 this information from one of the figures.

92 The term "noisy" has been replaced by "oscillating" which is more appropriate.

93 P2L30: Sentence starting with "An accurate representation ...". I do not see the link between
94 this sentence and the rest of the paragraph. Please rephrase.

95 The sentence is replaced by:

96 "Therefore, the change of the stratosphere in response to anthropogenic climate
97 forcing requires an accurate representation..."

98 P2L35: To my knowledge (http://www.geosci-model-dev-discuss.net/gmd-2016-40/), the
99 European Centre for Medium-Range Weather Forecasts (ECMWF) is having a version of their
100 Numerical Weather Prediction (NWP) system with a state-of-the art chemical equation
101 systems for the stratosphere (the same as BASCOE) as part of the Copernicus Atmosphere
102 Monitoring Service (CAMS). Even if the CAMS resolution is lower than the operational
103 version of ECMWF's NWP, the paragraph should be rewritten as the statement that
104 resolving state-of-the art chemical equation systems for the stratosphere is much too
105 expensive for NWP models is not completely true anymore.

106 P2L35-L44: "The computing time ... present-day conditions" is replaced by

107 "NWP models, Global Climate Model (GCM) and Earth-System Model (ESM) include
108 dynamics, chemistry, land-surface, sea-ice and ocean processes, each with some
109 degrees of complexity. For this reason, simplified chemical schemes are being
110 investigated (e.g. Baumgaertner et al., 2010) besides detailed chemistry modules
111 (e.g. Huijnen et al., 2016). Linearization of the stratospheric chemistry (Hsu and
112 Prather, 2010; Monge-Sanz et al., 2013) could be a computationally attractive
113 approach though their implementation would require both further studies as well as
114 a full reanalysis, or at least a chemical-consistent representation of the key gases of
115 interest in the stratosphere, which is, however, not available for present-day
116 conditions."

117 P2L44: Similar comment as before. I think CAMS operational real-time analysis and CAMS
118 reanalysis are also starting points for a full present-day stratospheric composition analysis
119 and reanalysis. This should be mentioned.

120 CAMS has the capabilities to assimilate stratospheric CH_4 and N_2O but this has not been
121 performed yet. As far as I understand, this is not planned (see Sect. 7.3 in Flemming et al.,
122 ACPD, 2016, http://www.atmos-chem-phys-discuss.net/acp-2016-666/)

123 P3L73: Maybe it worth mentioning in that paragraph if BASCOE accounts for the cross-
124 correlations between CH_4 and N_2O in the background errors.

125 BASCOE does not account for cross-correlations between CH_4 and N_2O in its **B** matrix, this is
126 not stated in the manuscript. P7L202 we replace "It has been calibrated..." by

127 "Background error correlations between species are not taken into account. Spatial
128 correlations have been calibrated ..."

129 P6L184: In the introduction it is claimed that NWP cannot afford a detailed stratospheric
130 chemistry (which is not completely right as discussed before) and that chemical transport
131 models are the alternative for that. But in this study, no stratospheric chemistry is
132 considered which would mean that BASCOE cannot afford a stratospheric chemistry too?
133 And the horizontal resolution of BASCOE for that study (3.75°x2.5°) is much coarser than the
134 CAMS operational system (TL511 or about 0.3°x°.3°). The author choice not to have a
135 stratospheric chemistry in this study should be more justified and harmonized with the
136 introduction.

137 The choice to turn off the chemistry is discussed in P6L183-L187. This will be stated in the
138 introduction, after P3L73 through two additional sentences:

139 "All the assimilation experiments performed in this study have been done with the
140 chemical scheme of BASCOE turned off. This is possible since the chemical processes
141 that affect CH_4 and N_2O occur on time scales which are much longer than the 24h
142 assimilation window used in this study."

143 P6L186: With an assimilation window of one day and no chemistry in BASCOE I would expect
144 to have model errors. Are these errors are accounted for in the assimilation (weak-constrain
145 algorithm)? If not, the authors should justify their choice.

146 No weak-constrain assimilation is used. Again, the chemistry can be turned off because the
147 chemical processes that affect CH_4 and N_2O occur on time scales which are much longer than
148 the assimilation window.

149 P8L227: If σ_o and σ_b are vectors, then they should appear in bold. The computation of σ_b
150 should then be detailed as it would be in the model space (37 levels) while σ_o would be in
151 the observation space. The square in the left-hand term should also be detailed. If the
152 square is an inner-product, then σ_o and σ_b are scalars and then we should know how they
153 are computed as they are errors on a retrieved profile and a model profile respectively.

154 In Eq. (4), σ_o and σ_b are scalars and σ_b is interpolated from the model grid to the observation
155 tangent point given by its latitude, longitude and pressure. Eq. (4) and the text that follows it
156 has been replaced by:

157 " $(\mathbf{y} - H(\mathbf{x}^b))_{ij}^2 > \alpha(\sigma_{ij}^o - \sigma_{ij}^b)$

158 where i and j denote, respectively, the number of the profile and the altitude index;
159 σ^o and σ^b denote, respectively, the error variance of the observations and of
160 background state, the latter being interpolated from the B matrix to the observation
161 location."

162 P8L258: I find the term "noisy" not specific enough. What the authors want us to see in this
163 figure? Despite the usage of the "noisy" term, I found this paragraph confusing: CTR is noisy,
164 BASELINE reduces the noise but the noisy structure in BASELINE is also present in BASEv7.
165 Maybe this could be clarified.

166 The term "noisy" has been replaced by "small structures" and the text between P8L257-L263
167 has been replaced by:

168 "In those conditions, no physical or chemical processes can explain the small
169 structures in the zonal mean of CH_4 in experiment RAW (e.g. local maxima around 80
170 hPa in the tropics or at 30 hPa and 50°S). These structures are smoothed when the
171 assimilation system uses the averaging kernels (exp. BASELINE) and a calibrated **B**

172   matrix (exp. ENS). The use of the vertical correlations in the **R** matrix apparently has a
173   very small impact when comparing experiments ENS and ENS-CR. BASEv7 shows
174   larger CH_4 volume mixing ratios than BASELINE, between 5 and 10 % depending on
175   the region. It is also found that the small structures in BASELINE are still present in
176   BASEv7."

177   P9L293: I do not see in Fig. 7 that ENS-CR has a high variability. I also disagreed with the
178   following statement: "This suggests that the observational error covariance matrices
179   provided by the MIPAS ML2PP retrieval are not optimal for data assimilation." This could
180   also suggest that the observational error variance used in BASCOE are not correct. For
181   example, one could inflate the observational error variance provided by the MIPAS ML2PP
182   retrieval to account for the representativity error. Increasing the observational error
183   variance could lead to reduce $J(x_a)$ (see general comment too).

184   I guess you meant Fig. 6? See our response to your general comment above and consider the
185   following changes:

186   P9L293 "...and with high variability" removed.

187   P9L293-L296 "This suggests that ..." replaced by:

188   "BASEv7 has values of $J(x_a)/p$ which are larger than those from BASELINE and further
189   larger than 1/2. ENS-CR shows a much larger deviation from theoretical values. This
190   suggests that the observational error covariance matrices provided by the MIPAS
191   ML2PP retrieval have not been used optimally by BASCOE. In particular, the
192   observational error variance of ML2PP retrieval probably needs to be tuned and
193   inflated to account for the additional weight of the observations due to their
194   correlation. This has not been done in this study."

195   P10L318: "...best to worse: ENS, ENS-CR, BASELINE ..." replaced by

196   "...best to worse: ENS, BASELINE ...".

197   P10L321-L324: "While providing ... further investigated." replaced by

198   "Experiment ENS-CR has been excluded from the ranking. As mentioned above, this
199   experiment need further test to take into account the additional weight of the
200   observations due to their vertical correlations."

201   P10L324: I think this should be investigated in particular by changing the observational error
202   variance in ENS-CR.

203   This has not been done, see our response to your specific comment above.

204   P11 paragraph starting line 353: I have some issues with the discussion of the differences
205   between MIPAS and ENS when MIPAS are assimilated and the title of the section is
206   "Validation Against Independent Observations".

207   I understand the confusion of the reviewer. The title of the section is renamed "Comparison
208   with Independent Observations".

209   P11L356 "The agreement between MIPAS and ENS is very good for CH_4 and N_2O" replaced
210   by

211   "As expected, MIPAS and ENS are in good agreement for CH_4 and N_2O".

212   3. Minor revisions/comments

213 P1L4: "The retrieved CH_4 and N_2O profiles" instead of "The CH_4 and N_2O profiles" to avoid
214 the confusion as it could be the profiles from BASCOE.

215 Corrected.

216 P1L9: "independent observations". Maybe you could provide the list of independent
217 observations (ACE-FTS for CH_4 and N_2O and MLS for N_2O).

218 P1L9: "... and independent observations demonstrating the general good quality of these
219 two retrievals provided by MIPAS ESA." replaced by

220 "... and independent observations from ACE-FTS (CH_4 and N_2O) and MLS (N_2O),
221 demonstrating the general good quality of CH_4 and N_2O retrievals provided by MIPAS
222 ESA."

223 P1L14: "CH_4 and N_2O observations". Strictly speaking, these are not observations but
224 retrieved data. Please be careful in the text with the usage of "observation".

225 "Observations" replaced by "retrievals". As well as:

226 P2L61: "two sets of observations" replaced by "two datasets"

227 P2L66, P5L155, P11L331, P13L413: "observations" replaced by "profiles"

228 P4L107: "vertical grid of the observations" replaced by

229 "vertical grid of the retrieved profiles"

230 P3L56: "CH_4 and N_2O are both emitted at the Earth's surface" instead of "CH_4 and N_2O are
231 both produced at the Earth's surface" as there is some (chemical) CH_4 and N_2O production
232 higher up in the atmosphere as well.

233 Corrected.

234 P3L60: "these retrievals". It is not clear "these" is referred to. Please detail.

235 P3L60 "these retrievals" replaced by "the retrievals of these two species".

236 P3L62: Do the issues are related to these particular retrievals or are they general issues?
237 Please precise.

238 The problem of the abrupt change in the gain affects only band B, and hence only the
239 species with spectral features in this band.

240 P3L63 inserted after "…addressed.":

241 "These issues are specific to the retrievals that use spectral points in band B, mainly
242 CH4 and N2O retrievals."

243 P3L66: "In their study" instead of "In that study"?

244 Corrected.

245 P5L142: What are "OR measurements"?

246 "OR" stands for "optimized resolution" as defined P3L86.

247 P10L 329: "MIPAS v7 does not improve" at all or significantly "the quality of v6 for CH_4 and
248 N_2O"? Please detail.

249 Some elements of the diagnostics used in this study show an improvement of MIPAS v7 with
250 respect to MIPAS v6 and other ones show a detorioration. In fact, no specific modifications
251 have been implemented in v7 to improve the quality of CH_4 and N_2O in the Optimized
252 Resolution phase (for the first phase of the project, the so-called Full Resolution, v7 uses
253 improved microwindows) . Therefore, we think that there is not a clear indication showing
254 that MIPAS v7 improves the quality of CH_4 and N_2O with respect to v6. We have modified
255 the sentence according to this conclusion. P10L329, after "in the tropics":

256   "According to the diagnostics that have been set up for this study, there is not a clear
257   indication that MIPAS v7 improves the quality of CH_4 and N_2O with respect to v6."

258 P11L339: sentence not clear. The ranges found by De Mazière et al. (2008) are the ranges of
259 ACE-FTS incertitude?

260 P11L339 "…, i.e. within the ranges found by De Mazière et al. (2008)." replaced by

261   "… . This is in the range of comparison between ACE-FTS CH_4 and independent
262   observations (De Mazière et al., 2008)"

263 P11L347: please give details on what Sheese et al. (2016) compared to have similar values.

264 This is detailed in Sect. 2.2. P11L347 "... found by Sheese et al. (2016)" replaced by

265   "those found in validation studies (see Sect. 2.2)"

266 P11L351: please provide values of these uncertainties.

267 They are given in Sect. 2.2 and 2.3. P11L356 "...each independent dataset" replaced by

268   "...each independent dataset (see Sect. 2.2 and 2.3)."

269 P11L353: please precise if MIPAS averaging kernels have been used to compute the figure.

270 P11L355 after "(N_2O only)":

271   "ENS and CTRL values are obtained using the averaging kernels of MIPAS."

272 P12L374: I am confused. MIPAS discontinuities are illustrated Fig. 10 but the next sentence
273 says that "the figure presents time series of daily averaged MLS N_2O", not MIPAS. I
274 understand what you mean but the paragraph should be rephrased.

275 P12L374: The sentence "Such discontinuities are relatively frequent in the whole period of
276 MIPAS optimal resolution as illustrated in Fig. 10." Is removed.

277 P12L386: you should detail the grid you used to compute these statistics (30° latitude band
278 on the horizontal, and in the vertical?).

279 P12L388 " ..., as a function of altitude and latitude and for the period October 2007-April
280 2012." replaced by

281   " ..., for the period October 2007-April 2012. These correlations are calculated on 30°
282   latitude bins (60° at the Equator) and on a pressure grid with 6 bins per decade of
283   pressure."

284 P13L414: you could detail **B** with "background error" here, just for the conclusion. Please
285 rephrase as not all experiments are using a **B** matrix calibrated using an ensemble method.
286 Moreover, you could also add that models (BASCOE for this study) are additional
287 information to the raw observations in assimilation systems.

This sentence is for an assimilation system with a calibrated B matrix so I did not rephrase it to mention that not all experiments are using a B matrix calibrated. I could add that a model was also used but it is implicitly part of the "assimilation system".

P13L414: "B matrix" replaced by "background error covariance matrix"

P13L432: could you please precise how this study shows that data assimilation can be considered a useful validation tool for geoscientific datasets?

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

> "Nevertheless, this study also diagnoses two issues in MIPAS CH4 and N2O 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 methods 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.

P21 Fig 5: why showing only the tropics? Maybe this could be discussed in the text. For the caption, I would use "scatter-plot" instead of "correlation" as the figure do not plot correlations. I would also add the labels (a) to (f): "observed by (a) ACE-FTS and from (b) to (f) for five BASCOE experiments".

The tropics are shown in order to provide a more compact scatter-plot than showing data from all latitudes.

P9L278 after "observation operator.":

> "Data is shown only for the 30°S-30°N latitude band in order to provide a more compact scatter-plot and better illustrate the different results obtained by each assimilation experiment."

P21 Caption of Fig. 5, P9L276, P9L280, P9L286, P10L327: "Correlation" replaced by "scatter-plot".

P21 Caption of Fig. 5: Labels (a) to (f) also inserted.

P25 Fig 9: it would be nice to have the labels (coloured line and legend) inside one of the subplots as for other figures.

Done