Authors' response to reviews of: "Comparison of the GOSAT TANSO-FTS TIR CH_4 volume mixing ratio vertical profiles with those measured by ACE-FTS, ESA MIPAS, IMK-IAA MIPAS, and 16 NDACC stations"

Kevin S. Olsen[1], Kimberly Strong[1], Kaley A. Walker[1,2], Chris D. Boone[2], Piera Raspollini[3], Johannes Plieninger[4], Whitney Bader[1,5], Stephanie Conway[1], Michel Grutter[6], James W. Hannigan[7], Frank Hase[4], Nicholas Jones[8], Martine de Mazière[9], Justus Notholt[10], Matthias Schneider[4], Dan Smale[11], Ralf Sussmann[4], and Naoko Saitoh[12]

[1]Department of Physics, University of Toronto, Toronto, Ontario, Canada
[2]Department of Chemistry, University of Waterloo, Waterloo, Ontario, Canada
[3]Istituto di Fisica Applicata "N. Carrara" (IFAC) del Consiglio Nazionale delle Ricerche (CNR), Florence, Italy
[4]Institut für Meteorologie und Klimaforschung, Karlsruhe Institute of Technology, Karlsruhe, Germany
[5]Institute of Astrophysics and Geophysics, University of Liège, Liège, Belgium
[6]Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Mexico City, Mexico
[7]Atmospheric Chemistry Division, National Center for Atmospheric Research, Boulder, CO, USA
[8]Centre for Atmospheric Chemistry, University of Wollongong, Wollongong, Australia
[9]Belgisch Instituut voor Ruimte-Aëronomie–Institut d'Aéronomie Spatiale de Belgique (IASB-BIRA), Brussels, Belgium
[10]Institute for Environmental Physics, University of Bremen, Bremen, Germany
[11]National Institute of Water and Atmospheric Research Ltd (NIWA), Lauder, New Zealand
[12]Center for Environmental Remote Sensing, Chiba University, Chiba, Japan

Correspondence to: K. S. Olsen (ksolsen@atmosp.physics.utoronto.ca)

1 Anonymous Referee #1

1.1 General comments

> This paper describes a comparison of CH_4 profiles retrieved from the GOSAT TANSO-FTS TIR with measurements by ACE-FTS, ESA MIPAS, IMK-IAA MIPAS, and NDACC. Although this manuscript presents results that would be of interest to readers of AMT, I found some of the authors' explanations difficult to follow. Therefore, some revisions are needed before it can be accepted for publication.

We thank the reviewer for their careful review. We will address each specific comment and try to clarify our manuscript as well as possible. Our aim was to provide a clear and concise description of our methods and focus on our results, which we also believe are of interest to the community.

1.2 Specific comments

1. p1, line14-15: "with and without smoothing" p9, line13: "To reduce biases caused by over-counting, when comparing TANSO-FTS to MIPAS, and by smoothing, when comparing TANSO-FTS to ACE-FTS,. . ." What is "smoothing" in this study? Please add a detailed description in Abstract and text to help the readers. Additionally, the authors should explain why they show correlation results based on both smoothed CH_4 profiles (Fig. 8) and unsmoothed CH_4 profiles (Fig. 9). What can we learn from this comparison?

Smoothing is a general term used to describe an operation that reduces the magnitude or frequency of fine-scale structure in a signal. When comparing two atmospheric remote sensing instruments with different vertical resolutions, the instrument with finer vertical resolution will have more fine-scale structure in its retrieved profiles. Likewise, an instruments whose retrieval has less of a dependence on the a priori may also have more fine-scale structure in its retrieved profiles. The instrument with lower vertical resolution or more dependence on the a priori will have retrieved vertical profiles that look "smoother." In order to compare the results of two instruments that are intrinsically different, we apply smoothing to that with finer resolution in order to account for these instrumental differences. Our objective is not to compare the results from two different instruments, but to ask, if one of our instruments had the same vertical resolution, information content, and dependence on the TANSO-FTS a priori, would the retrievals for that instrument agree with those from TANSO-FTS. The process, as described by Rodgers and Connor (2003), is standard practice when validating remote sensing vertical profiles of trace gas VMRs. The purpose of smoothing and the method used are described in Sect. 6.1. At the first mention of "smooth" in the abstract, the sentence has been extended to include a brief description of the purpose of smoothing. In the introduction, a few sentences were added to describe the need for smoothing and refer to Sect 6.1 for the method.

The reason we present results with and without smoothing is that a data user may not apply smoothing to the ACE-FTS, NDACC, or MIPAS results. The objective of validation is not necessarily to measure the magnitude of the differences between the two instruments' retrievals, but to do so in the context of the sensitivity and information content of the instrument being validated. These results may be of interest to data users, so they have been included.

2. p7, line13: "internal variability for each instrument" Due to insufficient description, I don't understand the meaning of "internal variability" in Sect. 3 and Fig. 1. Green lines (TANSO-FTS) in Fig. 1 show the difference between the GOSAT TANSO-FTS CH_4 retrievals and the a priori profiles. On the other hand, blue lines (MIPAS) are the difference between IMK-IAA MIPAS and ESA MIPAS. I don't understand how were the internal variabilities of ACE-FTS (p7, line25-33) and NDACC (p8, line9-15) evaluated. Does "the variability of NDACC data" mean the difference between NDACC CH_4 profile and TANSO-FTS CH_4 profile? In addition, can the authors explain the reason why they were compared in the same figure despite a different definition?

Several changes have been made to the internal variability section to address your concerns. This study is not central to the paper, but we feel that it is important to provide context for the main results. The internal variability is the difference between measurements within each instruments data set, loosely defined (too much so, we agree). Due to the different measurement techniques (especially the data acquisition rate of ACE-FTS compared to TANSO-FTS and MIPAS), a single method to estimate this variability was not used. Furthermore, there were other calculations of interest to us, for instance, the IMK and ESA data products for MIPAS allow us to directly compare different retrievals made from the same observations, something we cannot with ACE-FTS or NDACC. The measurements compared in this study are made at different times and locations, sampling different air-masses, and are subject to noise, and considering the internal variability of each instrument addresses the magnitude of the effects caused by these differences. Several changes were made to make the purpose of this study clear and we have tried to eliminate our usage of the term "internal variability" since we are presenting different measurements. Changes were also made to the caption of Fig. 1 to reflect this.

The reason each variability profile is placed on the same figure is because they are to be qualitatively compared and we see no reason to unnecessarily create extra figures and paper length.

3. p8, line20-27: "coincidence criteria" There is a lack of explanation why the coincident criteria were set as "within 12 hours and within 500km" for ACE-FTS and NDACC and set as "within 3 hours and within 300km" for the MIPAS data. For example, did the authors examine latitudinal and longitudinal dependence of TANSO-FTS data within 500km or 300km? I would show the spatial variations of TANSO-FTS CH_4 in the colocation circle at a particular height (the upper or middle troposphere). In addition, can the authors discuss the validity of their method by comparing the coincidences (e.g., statistics for match-upped data) in present study to those in the previous validation papers on the GOSAT data.

The coincidence criteria were used to try to optimize the number of coincidences with TANSO-FTS, increasing the small number of NDACC and ACE-FTS coincidences, and reducing the large number of MIPAS coincidences. Because MIPAS makes measurements much more frequently, we have the freedom to demand much tighter coincidences in space and time. At the beginning of Sect. 4, where "coincident criteria" is used, this is made explicitly clear with the statements: "Due the coverage and data collection rates of each instrument, different coincidence criteria were used. In the case of ACE-FTS, which only records two occultations per orbit, and NDACC stations, which are stationary, the objective of the coincidence criteria was to maximize the number of measurements used. Conversely, in the case of MIPAS, which makes frequent observations, the objective was to reduce the number of potential coincident measurements." Furthermore, we point out: "…for MIPAS–TANSO-FTS coincidences within 12 hours and 500km, we find approximately 180,000 coincidences per month."

The TANSO-FTS data are collected with a high frequency in a sweeping pattern along the satellite ground track. The high inclination of 98° provides a near-polar orbit. The result is high-density, near-global coverage, with more observations near the poles because the satellite ground tracks are more tightly spaced at higher latitudes. Reducing the spatial dependence of the coincidence criteria will have a different effect on each satellite. The impact will be larger on ACE-

FTS because its measurements over the tropics are very sparse, but for ACE-FTS we used the wider criteria. A way to avoid this difference in sample sizes between the tropics and poles is to use a degrees-latitude criteria. The result is that over the poles we are comparing measurements that are very close together, while those over the tropics may be separated by hundreds to thousands of kms, which would have a larger negative impact on our study.

The criteria used match other validation studies that use ACE-FTS, NDACC, and MIPAS data, and also the validation studies of CH_4 for each instrument. These are, however, far too numerous to list here. Coincidence criteria for primary CH_4 validation studies have been added to the text.

The reviewer asks whether we can compare coincidence statistics to previous validation papers on the GOSAT data. Only one previous validation paper on TANSO-FTS TIR data is in press: Saitoh et al. (2015) compares the CO_2 data product to aircraft measurements. They use very tight criteria (100 km and 2 hr) and consider only 140 coincident profiles. However, this study is not comparable with ours since the aircraft flights are in-situ measurements, rather than remote sensing, so a tighter coincident criteria is needed. The TANSO-FTS SWIR XCH_4 data product was validated by Inoue et al. (2014) that included the ACE-FTS instrument. However, they use climatological data, not coincidences.

4. p16, line31-34 “We also compared the differences shown in Fig. 10 to TANSO-FTS retrieval parameters: land or sea mask, sunglint flag, incident angle, both along the scan path and GOSAT track path, and observation mode (see Kuze et al., 2009). We found no biases in our coincident TANSO-FTS dataset related to any of these parameters, or whether the observation was made during night or day.” Can the authors show the features of the GOSAT TANSO-FTS biases related to land or sea mask and the other parameters in the previous section (or in Appendix)? It is not appropriate to discuss these important points without showing here.

The reviewer asks us to show the features of the GOSAT TANSO-FTS biases related to land or sea mask.” However, there were no biases to show. We understand the reviewers comment that it may be inappropriate to have mentioned investigating these quantities without explicitly showing the results. Our objective is to show due diligence on our part. We are not making strong statements about the effects of any of these quantities on the TANSO-FTS data, and that is not the purpose of this paper. We feel that it may be detrimental to our manuscript to include several more figures and significantly increase its length, while not significantly adding to the conclusions of our work. An appendix may be created showing the relationships of our results to the ancillary data in the GOSAT data files (eight additional figures) at the discretion of the editor.

1.3 Minor revisions

1. p4, line32: “the Halogen Occultation Experiment” —> “the Halogen Occultation Experiment (HALOE)”

change made, thank you

2. p7, line38: “the IMK-IAA data has” —> “the IMK-IAA data have”

changed “has” to “have” for the plural of “data”

3. p12, line23: “have a much smaller affect on” —> “have a much smaller effect on”?

this is certainly a case where the noun “effect” is correct

4. p15, line34: The Pearson correlation coefficient R2 of NDACC (0.9929) is different from that shown in Fig. 8.

thank you, this is a typo. The figure is correct, R^2 is computed during the generation of the figure. We also changed the order of the list to match the order of the panels of the figure

5. p19, line13: Please update information on Bader et al., 2016, ACPD.

updated the reference.

6. p19, line10: in reference list of Côté et al. (1998), “formulations” —> “formulation”

corrected the typo in the title

7. p20, line19: Please update information on Errera et al., 2016, AMTD.

updated the reference

8. p21, line33: in reference list of Picone et al. (2002), “1486” $\rightarrow$ “1468”

changed the page number from 1486 to 1468

9. p21, line11: in reference list of Raspollini et al. (2014), “Annal. Geophys.,” $\rightarrow$ “Ann. Geophys.,”

corrected the journal abbreviation

10. p28: a legend of Fig. 2, “NDDAC” $\rightarrow$ “NDACC”

corrected the typo. Similar to the next two items, axis labels indicating cardinal direction are missing from manuscript version downloaded from the AMTD website

11. p34: In Figs. 8 and 9, “x” of “y = mx+b” is not printed. In addition, “R” of “R2” is not printed.

during the technical review of this manuscript, this problem was mentioned. However, this issue did not exist on our copies of the submitted manuscript. We have downloaded copies from the AMTD website, and, sure enough, characters have been stripped from the figure. This is not an issue we can change, but will have a special note to the editors upon re-submission and ensure that the correct figure ends up in the compiled document. Thank you for pointing this out

12. p35: In Fig. 10, the unit of “Latitude” is not printed.

similar as for Fig. 9, the degree symbol appears in our submitted manuscript and figure files. We will discuss with the editors and ensure this symbol appears in a future version

2 List of changes

Page and line numbers refer to the AMT Discussions paper.

- p. 1, line 13: "vertical profiles are smoothed using the TANSO-FTS averaging kernels." changed to "and smoothing is applied to ACE-FTS, MIPAS, and NDACC vertical profiles. Smoothing in needed to account for differences between the vertical resolution of each instrument and differences in the dependence on a priori profiles. The smoothing operators use the TANSO-FTS a priori and averaging kernels in all cases."
- p. 1, line 15: changed "examine…" to "and examine…"
- p. 2 line 3: changed "…4 % …" to "…4 % ($\pm \sim 40\,\mathrm{ppbv}$)…"
- p. 2 line 5: changed "…investigated …" to "…investigate …"
- p. 2 line 7: changed "…equator …" to "…tropics …"
- p. 2 line 12: changed "…CO_2…" to "…carbon dioxide (CO_2)…"
- p. 2 line 15: added the sentences: "In this work we compare TANSO-FTS measurements with those made by similar instruments in order to validate its quality. Any biases in the data product need to be well understood for it to be used by other researchers, and their discovery may lead to improvements of future versions."
- p. 2 line 16: removed "…between 0.37 and 1.6 μm, each around 0.02 μm wide and chosen to avoid H_2O and O_2 absorption. TANSO-CAI…"
- p. 2 line 12: changed "…CH_4…" to "…carbon dioxide (CH_4)…"
- p. 2 line 24: changed "…(PEARL) in Eureka, Canada" to "(PEARL) at 80° N in Eureka, Canada (Batchelor et al., 2009)."
- p. 2 line 26: added reference (Kurylo and Zander, 2000) after "…(NDACC)."
- p. 2 line 28: removed paragraph break
- p. 2 line 29: removed comma after "…local"
- p. 2 line 33: changed "…made in coincidence …" to "…coincident …"
- p. 2 line 30: removed sentence beginning "Any biases…"
- p. 3 line 6: inserted the paragraph: "The question we are asking in this validation study is not *what is the magnitude of the difference between TANSO-FTS retrieved* CH_4 *vertical profiles*, but: *given the vertical resolution, information content, and a priori dependence of TANSO-FTS, would* CH_4 *vertical profile retrievals derived from another co-located*

instrument's measurements agree with those for TANSO-FTS? To answer this question a smoothing operator is applied to the vertical profiles of the instruments with finer vertical resolution (and therefore finer structure in the vertical profiles). This smoothing operator, described by Rodgers and Connor (2003), and presented in Sect. 6.1, uses the a priori profiles and averaging kernels from TANSO-FTS. However, results without smoothing are also presented here, as they will be of interest to data-users. provide an indication of the magnitude of these effects."

- p. 3 line 8: inserted "in order to account for the structure intrinsic to a finer-resolution instrument" after the bracketed clause

- p. 3 line 22: replaced "12900–13200 cm^{-1}, 5800–6400 cm^{-1}, 4800–5200 cm^{-1}, and 700–1800 cm^{-1}. The fourth band is in the TIR and is used ..." with "the TIR band is between 700–1800 cm^{-1} and is used ..."

- p. 3 line 27: changed "coverage" to "spatial coverage"

- p. 3 line 28: changed "...methodology..." to "...nonlinear maximum a posteriori method used for..."

- p. 3 line 30: added comma after "...surface temperature..."

- p. 3 line 32: added comma after reference to Saeki et al., (2003)

- p. 3 line 33: added reference (Saitoh et al., 2009)

- p. 4 line 8: removed "...a circular..."

- p. 4 line 8: removed "...with an inclination of 74° ..."

- p. 4 line 33: changed "...ACE-FTS v2.2 data..." to "...ACE-FTS v2.2 $\mathrm{CH_4}$ data..."

- p 4. line 10: removed "SCISAT also carries the ACE-Measurement of Aerosol Extinction in the Stratosphere and Troposphere Retrieved by Occultation (MAESTRO) instrument, a dual spectrophotometer with a wavelength range of 285–1030 nm and a spectral resolution of 1–2 nm"

- p 4. line 25: changed "...and their Global..." to "...with their Global..."

- p. 4 line 30: inserted comma after "...$\mathrm{CH_4}$ profiles..."

- p. 4 line 32: added "(HALOE)" after the acronym's definition.

- p. 5 line 5: added the sentence: "Waymark et al., (2013) found a slight reduction in $\mathrm{CH_4}$ VMR in the v3.0 data near 23 km, and a larger reduction of around 10% between 35–40 km."

- p. 5 line 7: removed "(inclination of 98°)"

- p. 5 line 9: inserted comma after "...cloud parameters..."

- p. 5 line 11: inserted comma after "…2004…"
- p. 5 line 12: changed "but…" to "but with…"
- p. 6 line 20: changed commas surrounding "below 25 km" to parentheses
- p. 6 line 30: changed "references, some information for each instrument" to "spectral range and resolution, and references."
- p. 6 line 7: changed "…dynamical nature of the Antarctic atmosphere" to "…characteristic atmospheric dynamics over Antarctica"
- p. 7 lines 12-16: The first paragraph of Section 3 was re-written to read: "To provide context for the VMR differences found when comparing each instrument to TANSO-FTS, shown in Sect. **??**, we have examined the variability of retrievals made for each instrument. We are interested in determining whether the mean differences found when comparing TANSO-FTS to another instrument are comparable to the differences found when comparing pairs of retrievals for a single instrument. Each pair of observations compared in this study are made at different times and locations and subject to instrument noise and analysis errors. Examining the variability within each data set provides an indication of the magnitude of these effects. Because the observation geometries and rates of spectral acquisition are different for each instrument, our internal comparisons differ for each instrument. For example, TANSO-FTS and MIPAS have a much higher data density than ACE-FTS, which only makes two sets of observations per orbit."
- p. 7 line 18: replaced the sentence beginning with "To examine the…" with "TANSO-FTS vertical profiles tend to be similar to their a priori and, therefore, to each other. To provide context for our validation results, we computed the magnitude of the mean differences between the TANSO-FTS retrievals and their a priori. This is indicative of the instrument sensitivity discussed in Sect. **??** and shows by how much the retrievals deviate from the a priori."
- p. 7 line 20: replaced "…3000 TANSO-FTS…" with …3000 randomly selected TANSO-FTS…"
- p. 7 line 25: replaced "…ACE-FTS sunset/sunrise measurement in a year…" with "…retrieved profile from an ACE-FTS sunset/sunrise (occultation direction)…"
- p. 7 line 26: replace "…the sunset/sunrise measurement acquired on…" with "…that from…"
- p. 7 line 27: added "(which are in different hemispheres)" to "…sunrise occultations,…"
- p. 7 line 27: changed "…acquisition was not made during an orbit" to "…acquisition was not recorded during a subsequent orbit."
- p. 7 line 30: added "VMR" before "difference"
- p. 7 line 30: removed "between pairs of VMRs"

- p. 7 line 35: changed “are” to: “were made”
- p. 7 line 35: removed “set of”
- p. 7 line 26: added the sentence “This provides an indication of the impact of different retrieval algorithms on retrieved profiles.”
- p. 7 line 38: changed “has” to “have”
- p. 8 line 9: changed “we considered the ...” to “we compared pairs of observations made at an NDACC site on the same day. We considered only...”
- p. 8 line 10: removed “...and found a subset of NDACC measurements for each site that were made on the same day.”
- p. 8 line 10: removed “then”
- p. 8 line 11: moved “...for each pair of measurements ...” to the beginning of the sentence
- p. 8 line 11: replaced “...made on the same day...” with “...on the standard NDACC retrieval grid...”
- p. 8 line 11: replaced “...multiple profiles” with “multiple coincidences in a day”
- p. 8 line 12: removed “the” and “all” from “...the differences are all found...”
- p. 8 line 13: changed “...dates with several measurements” to “...several measurements from the same day”
- p. 8 line 16: removed “Of the four data sets”
- p. 8 line 17: changed “and MIPAS...” to “that MIPAS...”
- p. 8 line 17: changed “while NDACC...” to “and that NDACC...”
- p. 8 line 17: replaced the sentence beginning “The magnitude of...” with “The magnitude of the internal variability of the data sets is between $\pm 2\,\mathrm{ppbv}$ (e.g., for NDACC and ACE-FTS in the upper troposphere) and $\pm 3\,\mathrm{ppbv}$, or around $2\,\%$ (e.g., for TANSO-FTS and the lower limits of ACE-FTS).”
- p. 8 line 21: replaced “In the case of ACE-FTS, which only records two occultations per orbit, and NDACC stations, which are stationary, the objective of the coincidence criteria was to maximize the number of measurements used. Conversely, in the case of MIPAS, which makes frequent observations, the objective was to reduce the number of potential coincident measurements.” with “ACE-FTS has an inclination of 74° and operates in solar occultation mode, recording only two occultations per orbit, predominantly at high latitudes; the NDACC sites are stationary; MIPAS makes frequent observations at all latitudes; and the spatial distribution of TANSO-FTS observations is enhanced by its cross-track observation mode. In the case of ACE-FTS and NDACC stations, the objective of the coincidence criteria was to maximize the number of measurements used. Conversely, in the case of MIPAS, the objective was to reduce the number of potential coincident measurements.”

- p. 8 line 26: inserted reference (Vincenty 1975)
- p. 8 line 28: the following paragraph was added: "The criteria used in this study are comparable to previous CH_4 validation studies. For example, de Mazière et al. (2008) used criteria of 24 hours and 1000 km when comparing ACE-FTS CH_4 to ground sites, and 6 hours and 300 km when comparing ACE-FTS to MIPAS. Payan et al. (2009) used criteria of 3 hours and 300 km when comparing MIPAS CH_4 to ground- and satellite-based spectrometers. Laeng et al. (2015) used criteria of 9 hours and 800 km when comparing MIPAS CH_4 to ACE-FTS, and 24 hours and 1000 km when comparing MIPAS to HALOE.
- p. 9 line 9: added reference to Holl et al. (2016) to the end of the sentence ending "... of the means."
- p. 9 line 19: changed "z score" to "standard score"
- p. 9 line 32: inserted comma after "... data set differ..."
- p. 9 line 33: changed "... with widths indicative of the..." to "... whose full-width at half maximum (FWHM) can be used to define the..."
- p. 10 line 6: changed "Each panel shows the mean from 30 retrievals, with each averaging kernel interpolated to a common pressure grid for that instrument" to "Each panel shows the mean from 30 retrievals. Vertical profiles of pressure associated with each retrieval's averaging kernel matrix are, in general, unique, so a common pressure grid was selected for each instrument and averaging kernels were interpolated prior to averaging."
- p. 10 line 10: replaced semi-colon with a period
- p. 10 line 13: inserted comma after "... role..."
- p. 10 line 22: changed "... ACE-FTS and MIPAS is close ..." to "... ACE-FTS and MIPAS, shown in Fig. **??**e, is close..."
- p. 10 line 23: removed the sentence "This is shown in Fig. 3e."
- p. 10 line 22: inserted comma after "... development..."
- p. 10 line 27: replaced the sentence beginning "The trace of the ..." with "The trace of the averaging kernel matrix gives the DOFs. For example, DOFs for retrievals made by TANSO-FTS, IMK-IAA MIPAS, ESA MIPAS, and NDACC from observations over the Arctic, above 60° N, are shown in Fig. **??**."
- p. 10 line 32: inserted the sentence "The trends visible are seasonal and are related to opacity and water vapour content."
- p. 11 line 22: removed parentheses around clause beginning "$\mathbf{x}_a$ and $\mathbf{A}$ ..."
- p. 11 line 24: changed "or the retrieval ..." to "or when the retrieval ..."

- p. 11 line 28: added reference to Sepúulveda et al. (2014).
- p. 11 line 29: changed "…common pressure grid" to "…common pressure grid, as opposed to an altitude grid."
- p. 11 line 30: inserted the sentence "Extrapolation is needed to ensure that the length of $\hat{\mathbf{x}}$ matches the dimensions of $\mathbf{A}$ in Eq. 1.
- p. 12 line 21: changed "…looking for…" to "…identifying and removing…"
- p. 12 line 22: removed "and then filtering these events"
- p. 12 line 23: changed "affect" to "effect"
- p. 13 line 20: changed "…VMR decrease differs…" to "…VMR decrease occurs differs…"
- p. 13 line 24: changed "…below 90 hPa" to "…below the
- p. 13 line 28: changed "Above 100 hPa…" to "Above the 100 hPa level…"
- p. 13 line 24: inserted the sentence "No zonal biases were observed in the unsmoothed data."
- p. 13 line 33: inserted comma after "…zonally…"
- p. 13 line 35: changed "This study reveals the actual differences one would expect when using the TANSO-FTS data product" to "Fig. 6 shows the mean differences between the TANSO-FTS data product and those of other instruments"
- p. 13 line 37: changed "the differences…" to "the difference profiles in Fig. 6…"
- p. 14 line 21: changed "For consistency, partial columns…" top "For consistency, each pair of partial columns…"
- p. 14 line 23: changed "…partial columns" to "…for each coincident pair of profiles"
- p. 14 line 37: inserted the sentence "These excluded data do not exhibit a broader distribution, but their computed partial columns are all very small due to the integration range."
- p. 15 line 24: changed "0.9986, 0.9968, 0.9965, and 0.9929 for ACE-FTS, ESA MIPAS, IMK-IAA, MIPAS, and NDACC" to "0.9986, 0.9965, 0.9968, and 0.9958 for ACE-FTS, IMK-IAA MIPAS, ESA MIPAS and NDACC"
- p. 16 line 26: the sentence beginning "A bias is seen…" has been replaced with "Weighted least squares regression of the combined data sets for each hemisphere reveals a bias at all latitudes of $13.30 \pm 0.06\,\mathrm{ppbv}$."
- p. 16 lines 26, 28, 29: changed units from ppmv to ppbv
- p. 16 line 28: changed "…combined data set…" to "…combined data sets in each hemisphere…"
- p. 16 line 29: changed "…or…" to "…and…"

- p. 16 line 30: added the sentence "The biases are latitude-dependent and vary between the tropics and the poles."
- p. 16 line 32: inserted the sentence "Each parameter was compared to the latitudes and the mean differences in Fig. 10, and the regression and covariance statistics from least squares fitting were computed."
- p. 16 line 38: added the following paragraphs to the discussion:

The primary driver of the mean differences found when comparing TANSO-FTS to other FTS instruments, with and without smoothing, is the instrument design and observation geometry. TANSO-FTS is a much more compact and, therefore, coarser spectral resolution FTS than those used in the comparison. The coarser spectral resolution makes it harder to distinguish closely spaced absorption lines, leading to poorer vertical sensitivity and higher uncertainty in the measurements. While the TIR spectral range of TANSO-FTS is comparable to that of MIPAS, the mid-infrared ranges of NDACC and ACE-FTS include a very strong methane absorption band near 3000 cm^{-1} with little interference from $\mathrm{CO_2}$, increasing their sensitivity and ability to accurately constrain $\mathrm{CH_4}$ retrievals. Furthermore, MIPAS and ACE-FTS observe the limb of the atmosphere, providing them with more measurements per retrieved profile, improved vertical resolution, and much higher sensitivity. While NDACC instruments also only have a single spectrum per retrieved profile, they observe the sun directly (as does ACE-FTS), resulting in a very strong signal. All these factors contribute to TANSO-FTS performing retrievals on a lower spectral resolution measurement of a weaker signal compared to MIPAS, ACE-FTS and the NDACC sites. This results in the sensitivity and DOFs shown in Figs. 3 and 4.

In Sect. 3, we examined the variability within each data set. This gives an idea of some of the sources of error in our comparison. The coincidence criteria used allow for the comparison of retrieved $\mathrm{CH_4}$ vertical profiles from different air masses. Our investigation of the NDACC data provides an estimate of the dependence of the $\mathrm{CH_4}$ abundance on time, since we compared profiles retrieved from the same location using the same retrieval algorithms, but at different times of day. Our result shows that temporal spacing may contribute around 5 ppbv. Our investigation of the ACE-FTS variability fixed the instrument and retrieval algorithm, but compared observations of different air masses, and we found a similar result of only several ppbv. The largest variability was exhibited when we investigated the MIPAS data set. This comparison was of the same observations analyzed by different retrieval algorithms (IMK-IAA and ESA), and resulted in much larger mean differences on the order of 100 ppbv.

Differences in retrieval algorithms between TANSO-FTS and the validation instruments may also account for the differences found in Figs. 5 and 6. Small differences in spectroscopic parameters exist, for example, each instrument's retrieval algorithms use different editions of the HITRAN line list. Comparisons of these line lists, and their impact on retrievals, can be found in **???**. The most significant parameter for TANSO-FTS is its a priori due to the weight given to the a priori profile by the TANSO-FTS averaging kernels in the retrieval. In Sect. 3 we compared the TANSO-FTS retrieved vertical profiles of $\mathrm{CH_4}$ to the corresponding a priori profile and found that they differ, on average, by up to 30 ppbv. This provides a rough minimum of the accuracy of the a priori profiles required for the the retrievals.

- p. 17 line 10: Inserted the sentences “The TANSO-FTS TIR CH_4 vertical profile data product is an important and novel data set. Its vertical range extends lower into the troposphere than other satellite data products, and its spatial coverage is global with a high density of measurements.”
- p. 17 line 10: removed comma after “...the sensitivity”
- p. 17 line 10: removed “...CH_4 TIR vertical profile...”
- p. 17 line 12: removed “useful”
- p. 17 line 12: changed ‘below” to “around”
- p. 17 line 13: replaced the sentence beginning “However,...” with “Unfortunately, the lower altitude boundaries of the other satellite-based data products, between 7–15 km, reduces the vertical range over which we can make comparisons.”
- p. 17 line 14: changed “in the troposphere” to “in the upper troposphere”
- p. 17 line 19: replaced the sentence beginning “We found that...” with “We found that the shapes of the TANSO-FTS CH_4 VMR vertical profiles near 15 km, where the CH_4 VMR falls off with increasing altitude, does not match those of the other instruments, and in a consistent manner, resulting in a pronounced feature in the mean difference profiles in Fig. 5, just below the 100 hPa level.”
- p. 17 line 25: changed “...dependence of the mean differences, taken over altitude and latitude” to “...dependence of the vertically-averaged differences on latitude.”
- p. 17 line 26: Inserted paragraph break before the sentence starting “We look...”
- p. 17 line 27: Inserted the sentences “ In a future release, the a priori will not be changed, but remain the outputs of the NIES-TM. Kuze et al. (2016) used theoretical simulations to determine that the Level 1B spectra which were used (V161) to generate the current TIR CH_4 data product had considerable uncertainties. New Level 1B spectra are due for release in 2018 and should lead to improved retrievals. Kuze et al. (2016) also proposed some corrections to the TANSO-FTS TIR L1B spectra which may be implemented. The spectral line list used (HITRAN 2008) will be updated. Uncertainties in the surface emissivity over cold surfaces (snow and ice) affect the retrieval at higher altitudes and will be improved in the next release. Improvements are also being made to the way the retrieval handles and simultaneously retrieves interfering species, such as O_3.”
- p. 19 line 10: changed “formulations” to “formulation”
- p. 19 line 13: updated the Bader et al. (2016) reference to reflect a change from ACPD to ACP

- p. 20 line 19: updated the Errera et al. (2016) reference to reflect a change from AMTD to AMT
- p. 21 line 11: changed “Annal.” to “Ann.”
- p. 21 line 33: changed “1486” to “1468”
- Table 1: in footnote [a] changed “. . . are often. . . ” to “. . . may be. . . ”
- Table 1: added the footnote “The Altzomoni site came online in late 2012.”
- Table 1: added the footnote “The Maïdo, La Réunion site came online in early 2013.”
- Fig. 1 caption changed to read: “Results for investigating the variability within each CH_4 VMR profile data set. Shown are the following comparisons: TANSO-FTS retrievals compared to their a priori (green), pairs of sequential ACE-FTS retrievals (red), ESA MIPAS retrievals to IMK-IAA MIPAS retrievals made for the same limb observations (blue), and pairs of NDACC retrievals made on the same day (orange). All retrieved profiles used are coincident with TANSO-FTS. Dashed lines are one standard deviation.”
- Fig. 2 changed “NDDAC” to “NDACC” in the legend