Authors' response to: "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” → “1468”

changed the page number from 1486 to 1468

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

corrected the journal abbreviation

10. p28: a legend of Fig. 2, “NDDAC” → “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 Anonymous Referee #2

2.1 General comments

Olsen et al. (2017) compares GOSAT TANSO-FTS profile retrievals and partial columns of methane (CH4) to coincident data from ACE-FTS, ESA MIPAS, IMK-IAA MIPAS, and NDACC, focusing on the upper troposphere/lower stratosphere (UTLS) over 2009-2013. This work expands an earlier TANSO inter-comparison in the Arctic to include global measurements from additional satellite and ground-based remote sensing instruments, aiming to identify possible zonal variability in the profile retrievals. Given the importance of understanding global CH4 and the uncertainty in its recent trends, precise and accurate measurements with global coverage are very much needed. As GOSAT is one of the primary satellites in monitoring atmospheric CH4 today, the as- sessment of the spatial biases of TANSO measurements would give confidence to the scientific community's use of an important data set. The approach described in the paper is comprehensive, and the methodology is thought-through. However, the exposition of the significance, conclusions, and limitations of this work require more development. The paper would benefit from a rebalancing of its structure, with the most critical changes being: (1) augmenting the discussion of previous and/or similar validation efforts of TANSO CH4, (2) paring down of the instrument background sections to focus on details directly relevant to their results, and (3) discussing the reasons for and implications of the differences highlighted in the comparison. Therefore, I would recommend publication of this manuscript after these points are addressed.

We thank the reviewer for taking time to make such a thorough review of this manuscript. Their comments are about the general structure and language of the paper, and cover specific technical and grammatical points. We have worked very hard to address each comment with care and believe that the outcome is a more streamlined paper that better addresses the scientific merit of our study without distracting the reader with overly technical discussions. The three primary points made above are discussed in more detail below by the reviewer, asking pointed questions and offering examples. After careful consideration, we agree with the reviewer and have made several changes and additions to the manuscript,

2.2 Scientific evaluation general comments

1. This paper would benefit from a more rigorous analysis (or if already done, a more comprehensive description) of the causes of the differences noted in the results. The authors are thorough in considering different parameters, but the text lacks a synthesis of how these difference relate to the results of the comparisons. For example, what component of the differences between TANSO and the other instruments can be explained by spectroscopy verses a priori profiles used? This can be addressed for the specific parameters already mentioned in the text; e.g. many of the retrievals incorporate the same linelists–do those profiles show better agreement to each other? The implications of these differences for the scientific community's use of the TANSO product based on these differences would also strengthen the paper.

Some of the primary comments from reviewer #1 are relevant to Sect. 3 where we look at the variability in each dataset. A discussion of these results in the context of our inter-instrumental comparison is lacking. We agree that the text lacks a synthesis and have made significant changes to Sect. 3 and added text to our discussion in Sect. 8. We think this reasonably addresses the reviewers comments and strengthens our manuscript.

Some of the specific questions raised in this reviewer's comment are, we feel, beyond the scope of this research topic and manuscript. Such aspects as the TANSO-FTS retrieval's dependence on the a priori or line list can be insinuated from our work, but not properly quantified. We have made a better effort to comment on our results in Sect. 8, but, since we cannot vary these parameters, redo the retrievals, and test the effects of these parameters, thoroughly addressing these questions is best left to a future, specific manuscript about the TANSO-FTS TIR CH_4 vertical profile retrievals.

The impact of differences in spectroscopy and retrieval methods on our results is difficult to quantify, but these factors certainly make significant contributions. For example, the retrievals for the instruments we considered use different versions of the HITRAN line list. There are significant and much discussed differences between CH_4 line lists in each version and CH_4 spectroscopy may be a limiting factor in CH_4 retrieval accuracy. The next version of the TANSO-FTS TIR CH_4 data product may use a different version of HITRAN and the TANSO-FTS team will quantify the impact of the line lists.

We quantified the difference between the IMK-IAA and ESA MIPAS retrievals by comparing pairs of profiles retrieved from the same MIPAS observations. This shows the impact of the retrieval algorithm and spectral line lists. The discussion of this in the manuscript has been improved to highlight this. The largest effect on our study, however, is due to physics and instrument design. TANSO-FTS is a much lower resolution instrument than MIPAS, ACE-FTS and those used by NDACC. Viewing geometry is what led to the use of a retrieval algorithm that incorporates the a priori into its solution, and also results in smaller signal strength than the other instruments. Our discussion has been extended to more thoroughly discuss our findings in Sect. 3, and to address the differences in instrument capability and retrieval algorithms.

2. At several points throughout the paper, principally in Sections 2-5, the authors include technical details about the satellites/instruments that do not seem pertinent to the study presented (e.g. the inclination angles). The extraneous information was distracting and diluted the narrative of the instrument comparison. I recommend including only those details that the reader should know to understand and evaluate the results and conclusions of this comparison, especially if those details are published elsewhere. For necessary details where the connection to the present study is not clear, explicitly listing the relevance and/or implications would be useful. (e.g. The authors list measurement windows and spectral resolution but need to comment on how these inter-instrument differences might relate to the results.) The authors might also consider moving some of these details that are useful but not central to the paper into an appendix or supplement.

While we believe these details are necessary for the reader, we understand the drawbacks pointed out by the reviewer, especially thanks to several specific examples. We removed the ranges of the non-TIR TANSO-FTS channels, the details about MAESTRO on SCISAT, details about TANSO CAI, and the inclination of MIPAS. The inclination of SCISAT is important, since it affects the location of solar occultation measurements which leads to far more coincidences at high latitudes. The inclination has been moved to the coincidence discussion, Sect. 4, and follows a statement about the distribution of the ACE-FTS measurements. In the discussion, Sect. 8, we have addressed the different spectral ranges and resolutions, and their possible impact on our study.

3. Given that methane is generally provided in units of ppb, do the authors have a specific reason for using ppm? If not, I suggest changing the references and figures to ppb. Figure 5 in particular would be more clear without extraneous zeros and decimal places.

When discussing methane on Earth, the abundance in the troposphere is around 1.5–1.7 ppmv, and the literature commonly uses ppmv when discussing CH_4. Our work is consistent with other, related literature sources, such as the CH_4 validation studies for MIPAS and ACE-FTS (Payan et al., 2009; Pleininger et al., 2015; de Maziere et al., 2008) and the preceding work by Holl et al. (2015). That being said, we do agree with the reviewer about Fig. 5 and have changed its x-axis from ppmv to ppbv. We have left Figs. 1 and 6, since we would be adding zeros to the x-axes of those graphs. Furthermore, we agree that sticking to ppmv when discussing CH_4 *differences* (e.g., Sect. 8) is cumbersome, and we have made several changes from ppmv to ppbv in the text.

2.3 Specific comments

1. p. 2 l. 11-13: The redundancy in the list of GOSAT objectives can be pared down. In addition, these objectives should be related, at least in part, to the objectives of this research. i.e. How does this paper contribute to the objectives of the satellites and the scientific community? This question is briefly touched upon on l. 31, but needs to be developed.

We have broken the short paragraph at line 29 up, placing the argument mentioned at line 31 with the GOSAT objectives at lines 10-13.

2. p. 2 l. 24-26: Given the focus on zonal dependence, providing the latitude of Eureka in the text would be useful. Also, please include citations for PEARL and NDACC.

We provided the altitude of Eureka and added two citations.

3. p. 2 l. 29-31: As the main objective (and contribution) of this paper is to expand TANSO validation globally, more consideration of the issues of spatial coverage is needed, including a literature review of zonal biases (possibly does not exist for this particular TANSO product, but if that is the case this should be stated) and a description of mechanisms that make the Arctic non-representative (e.g. polar vortex changing vertical profiles of trace gases and reducing the accuracy of a priori information).

This paragraph was removed as part of the correction for comment 1. Such a study for the TANSO-FTS data was performed by (Holl et al., 2016) and his results are discussed in Sect. 6.1.

4. p. 3 l. 5/Table 1: Please add year ranges to Table 1. It is not clear when mentioning the 2009-2013 time frame in the abstract whether all of the instrument measure consistently over that time period.

This is not so easy to address, all the instruments have downtime, and those operated at high latitudes are run in seasonal campaigns. Only two of the instruments came online during our study, Altzomoni and La Reunion Maido, and those dates have been added as footnotes to Table 1.

5. p. 3 l. 7-14: If this is an outline, put section numbers after each sentence. If this is an overview, this paragraph might fit better in the methods section rather than the introduction.

Appropriate section numbers are given on lines 6-7 and 15-19.

6. p. 3 l. 27: What type of coverage? Spatial? Spectral?

This is spatial coverage, added the qualifier to the text.

7. p. 3 l. 28: Is this paper the first time methodology for CO2 is applied to CH4? If so, the retrieval (or the at least the aspects that differentiate it from other TANSO retrievals) should be described more.

We added more details about the algorithm used, which is more in-line with our presentation of the MIPAS and NDACC data sets.

8. p. 3 l. 33: Do the 2011 updates have a reference (e.g. on the HITRAN website or used in a validation paper)? If so, please include a citation.

There is not a citation for this. The HITRAN 2008 paper refers users to the HITRAN website.

9. p. 4 l. 4-6: This sentence, with the important conclusions of the referenced paper added, should be moved to the Introduction, at the end of the second paragraph.

This is a summary of what is already in the introduction. Page 2, lines 22–28 discuss this in greater detail and provide the actual, quantitative conclusions.

10. p. 4 l. 9-11: If information about MAESTRO is relevant, include a reference for the instrument; otherwise this sentence can be removed.

Technical details about MAESTRO have been removed.

11. p. 4 l. 16: Is 5km the lowest altitude for ACE-FTS measurements filtered using the recommended flags (e.g. not a priori values)? Listing the lowest altitude of the data used in this paper would be more relevant, particularly for the discussion on vertical range in subsequent sections.

5 km is approximate, some occultations may extend lower. As can be seen in Figs. 5 and 6, our data extend to 6 km.

12. p. 4 l. 33 and 36: Do the percentages listed apply to all trace gases or just methane? Please make this more clear.

Added specific references to CH_4. The validation paper de Maziere et al., (2008) focuses only on CH_4.

13. p. 5 l. 4-5: Because this is the data version used, the results from Waymark et al. (2013) should be summarized, as is done with the above papers.

Waymark et al., (2013) contains a table summarizing the results fro each gas. The result for CH_4 was added to the manuscript.

14. p. 5 l. 25: By "Initial guesses" do you mean a priori profiles? Please clarify.

The ESA MIPAS team makes a distinction between “initial guess” and a priori. The initial guess is a generic profile used to begin the least squares fitting iterations. The term a priori means that we have prior knowledge of the state of the atmosphere and that the initial profile has a level of accuracy that the “initial guess” may not.

15. p. 6 p. 5l. 6-7: What are the reasons for the outliers/discontinuities? Would these impacts the results of this study?

Outliers would hopefully have little effect on this study due the large sample size we use. We rely on the ESA MIPAS data quality flags to remove spurious results.

16. p. 6 l. 30: Using "some information" is vague and does not tell the reader the relevance of the included information.

We agree, “some information” is vague. That may have been written before we decided on what information to include in Table 1. The specific columns are now in place of the vague terminology.

17. p. 6 l. 32-35: Is there a reference for an inter-comparison of these instruments? How are inter-site differences due to different instrumentation accounted for in this study?

There is not a publication detailing exactly what the reviewer asks yet, but the reader is invited to see Table 1 for publications describing each site.

18. p. 6 l. 36: Are those references for the most recent versions of the retrieval software? The spectroscopy has presumably changed since 1995/2004.

These references are not for the most recent versions of the software, but are the most recent publications about the software.

19. p. 6 l. 37: Does “harmonized” mean consistent between sites? Please use a more clear term.

Yes, harmonized means to “make consistent or compatible,” so in our context we mean that the site operators have attempted to make the retrievals consistent between sites. This is the terminology used by the NDACC InfraRed Working Group.

20. p. 7 l. 10: Given the differences delineated in this section, have you done a covariance analysis or sensitivity test to assess whether the results of the comparison depend on the retrieval software used, instrument, or any other difference across NDACC sites?

Please see line 11, the start of Sect. 3. This comment is addressed as part of the reviewers Scientific Evaluation.

21. p. 7 l. 20: Are these measurements representative of year, season, location, tropopause height, etc.?

These data were randomly selected from our sample to avoid diurnal, seasonal, or zonal effects. This information has been added to the text.

22. p. 7 l. 25: Please define "sunset/sunrise measurement" in this context.

Noted that sunset and sunrise refer to the occultation direction. In solar occultation geometry, the satellite instrument makes a series of observations by directly viewing the sun while the limb of atmosphere intersects the line of sight. In each orbit there are two occultation opportunities: when the satellite enters the shadow of the Earth, and when the satellite exits the shadow of the Earth. The instrument observes sunsets and sunrises, respectively.

23. p. 8 l. 5-6: Please summarize the results that the bias is consistent with.

This is in Sect. 2.3.2 where Laeng et al. (2015) is summarized.

24. p. 8 l. 24: Why was reducing coincident measurements an objective, if you could average them and thereby reduce potential bias (c.f. Kulawik et al. 2016)? Was this a data processing issue from the large number of coincidences? If the coincidence criteria varies by instrument, some sort of bootstrap or sensitivity test with a subset of data should be run to see if the VMRs are different with the more lax coincident criteria.

For consistency, we wanted to have sample sizes for the MIPAS comparison to be of the same order of magnitude as those for NDACC and ACE-FTS.

25. p. 9 l. 8-9: Including references for the approach mentioned would be useful.

Added a reference to Holl et al. (2016) for example.

26. p. 9 l. 19: Please define a z score and/or include a reference.

z-score is a standard parameter in statistics found in most undergraduate textbooks. It is perhaps more appropriately referred to as the “standard score,” and it has been changed.

27. p. 9 l. 26-27/Figure 2: Does using only the first 200 observations of the year capture any time-varying spatial coverage of the satellite data? Would it be more appropriate to use the first 20 observations of each month of 2012, for instance?

The orbit actually repeats its flyover-locations every couple of weeks. Using the first twenty occultations from each month would introduce a bias. However, the figure is purely qualitative. It shows the NDACC sites, that ACE-FTS favours high latitudes, and that there are no ESA MIPAS coincidences over the tropics.

28. p. 10 l. 5-6/Figure 3: Please provide a more clear description of what the pressure levels in the legends correspond to (i.e. the pressure widths of the averaging kernel rows vs. the pressure on the y-axis), as relates to the findings of the paper. (You could perhaps include the simple averaging kernel equation if useful, but if the text becomes too detailed I would suggest moving this description to an appendix/supplement.) Also, if these pressure levels are meant to be compared across instruments, using a single colour scale for panels a-d would be helpful, e.g. binned into ranges or following a colour gradient.

Since pressure is a retrieved parameter for some of the instruments, each retrieval is done on a unique pressure grid. To compute means of the averaging kernels, each kernel has to have the same vector length corresponding to a common pressure grid. Since Fig. 3 is illustrative, the pressure grid for each panel is not important, but that the means are taken at the same pressure level is. The pressure levels of the averaging kernels are given in each data product. A sentence was added to clarify this. The colours indicating each instrument are consistent throughout the manuscript.

29. p. 10 l. 8-12: Please add the implications for these comparisons. Related to the previous comment, the remarks on "full-width at the half-maximum values" would be more understandable by rewording the phrase "values when considering the location of the appropriate pressure level" and adding a more clear description of the averaging kernel widths (p. 9 l. 33-34).

The implication of these findings is given on line 8, that TANSO-FTS is the instrument with coarser vertical resolution, and on line 25 after further description of the figure, that TANSO-FTS retrievals are highly dependent on their a priori. We've re-organized this paragraph to be more clear, introduced FWHM notation, and moved the definition of vertical resolution to the discussion of kernel widths.

30. p. 10 l. 25-26: How do you determine the influence of this dependence on the results? (This paragraph might need to be moved closer to the discussion of TANSO priors later in the manuscript.)

This paragraph is a description of Fig. 3 and follows the preceding discussion of MIPAS and ACE-FTS sensitivity naturally. The implication is explored further later in the manuscript, especially in the discussion, Sect. 8, which was greatly expanded to address comments from reviewer #1.

31. p. 10 l. 32: What is the implication of the flat trends over mid-latitudes and tropics, as relates to the objective of this paper to determine zonal dependence of the retrievals?

The trends are related to seasons, their southern hemisphere reversal is actually the phase difference of the seasons. This has been re-written to reflect that.

32. p. 11 l. 28: Please include a reference for the claim that the NDACC a priori/ measured pressure profiles are accurate. (This might fit better in Section 2.)

Added a reference to the NDACC retrievals paper (Sepúulveda et al., 2014), which quantifies the contributions of sources of error, including NCEP a priori.

33. p. 11 l. 27-29: I'm not sure I follow the logic here; does this just affirm that interpolating to a common pressure grid does not introduce additional bias or uncertainty?

What we are defending here is not the choice of pressure grid, but the choice of doing our analysis in terms of pressure levels instead of altitude. We argue that pressure is more well known than altitude. We have added a statement specifying this.

34. p. 11 l. 30-32: Do these extrapolated values actually become part of the profile comparison? If not (as indicated in step 5 on p. 12 l. 10-11), why extrapolate these values at all? If so, profiles with extrapolated values included in the comparison would be problematic: the minimum altitudes for these instruments are so high that they tend to be in the region of the atmosphere where CH4 varies significantly with altitude, and the extrapolation would have large uncertainties.

A common vertical grid is vital since we are finding the means of profiles. Extrapolation is needed because the vector lengths need to match those of TANSO-FTS averaging kernels. Extrapolated pressure levels are not used. A statement was added that explains the need for extrapolation and refers to Eq. 1.

35. p. 11 l. 32-33: Is this sentence a way of saying that the averaging kernel equals zero where no measurements exist? Wouldn't this zero out the extrapolation referred to in the preceding sentence?

In the non-overlapping region, we have set values of $\hat{\mathbf{x}}$ equal to $\mathbf{x}_a$. When $\mathbf{A}$ is multiplied by $(\hat{\mathbf{x}}-\mathbf{x}_a)$, $(\hat{\mathbf{x}}-\mathbf{x}_a)$ nullifies the contributions from the averaging kernel matrix at pressure levels where $(\hat{\mathbf{x}}-\mathbf{x}_a)$ is zero.

36. p. 12 l. 16-17: Do you look at the longitudinal variability for each zonal band? If so, does it vary between instruments?

The longitudes of coincidences are fairly uniformly distributed in each zone for ACE-FTS and MIPAS, but are not for NDACC.

37. p. 12 l. 17-20: If my understanding is correct that you apply these filters only for individual points, as opposed to the entire profile, how do you account for heterogeneity in the underlying profile? For example, are all seasons represented at most altitudes ranges? If representation bias is not accounted for, the differences between the underlying measurements might account for some of the features illustrated in Figure 5, e.g. the strong agreement at high vs. low altitudes.

We use the data products as instructed by the product documentation, in some cases, specific points, usually at altitude extrema, are discarded, in others, the entire profile is discarded (e.g., when we say the ACE-FTS *quality_flag* cannot be equal to 4 at any altitude). Seasonal effects are smaller than zonal effects, the strongest being the increased cloud and humidity over the tropics, which strongly reduces the number of overlapping coincidences from MIPAS. Seasonal effects are shown in Fig. 4, and the strongest effect is for IMK MIPAS. The decrease in IMK MIPAS DOFs in Arctic winter corresponds to fewer measurements in the observation (ESA MIPAS, TANSO-FTS, and ACE-FTS do not have such a seasonal dependence). The features the reviewer mentions in Fig. 5 are attributed to the TANSO-FTS averaging kernels tending to zero at higher altitudes. Where the comparison is in very good agreement, the TANSO-FTS is approximately equal to its a priori, and the smoothed target profile has been transformed into approximately the a priori.

38. p. 12 l. 21-22: I find it surprising that measurements within the polar vortex did not impact the results, unless the problematic profiles were discarded through other filters or a priori values were used. Do Holl et al. (2016) apply the same data flags listed in this paper?

Holl et al. (2016) did use the same data quality flags, those for for ACE-FTS are referred to in their manuscript. The primary reason the effect is negligible here is that the temporal duration of the polar vortex is small in our data set.

39. p. 12 l. 28-29: Following on comments made on the manuscript's introduction, an explanation of why zonal biases may exist should be included. l. 29-31/Figure 5: Reiterating the last general comment, units of ppbv in the left-most panel would remove some of the extra text on the axes associated with the decimals and might be more intuitive for CH4.

Statements explaining this were added to the introduction, where Holl et al. (2016) is first discussed.

40. p. 12 l. 31-32: Do the sizes of the bins alter the comparisons? Are these zonal ranges narrow enough in the Northern Hemisphere? i.e. Are the profile differences at 50-60N comparable to 30-40N?

The x-axes of the left-most panels of Fig. 5 have been changed from ppmv to ppbv.

41. p. 12 l. 37: Please give a brief explanation of what this statistic tells us (or justification for using it) as opposed to the general correlation coefficient and/or include a citation.

Yes, the bins are narrow enough. Shrinking the zonal bin width results in a paucity of data over the tropics, and we tested different bin sizes and found no significant biases in any zone.

42. p. 13 l. 18-22: The pressure level at which CH4 decreases is the tropopause height. Unless I am misunderstanding this paragraph, the implication is that the tropopause heights of the instruments are different, which would very likely account for at least some of the profile differences observed. How do the calculated the tropopause heights compare among the various instruments? If they differ, it would indicate that some of the assumptions underlying the pressure interpolation (outlined in the paragraph on p. 11

The Pearson correlation coefficient is the most common form of correlation coefficient and most appropriate for this type of study. The proper name distinguishes it from rank, distance, weighted, adjusted, and other forms of correlation coefficient. What the reviewer refers to as a "general correlation coefficient" is most likely the Pearson correlation coefficient.

43. p. 13 l. 25-29) might need to be reconsidered. Measurements for which the a priori values have a significant influence could be especially susceptible to tropopause height biases.

We believe the observed phenomenon does not represent the physics of the atmosphere. Since the TANSO-FTS averaging kernels fall off at around the tropopause height, its retrievals have low sensitivity at that altitude. The TANSO-FTS a priori may not accurately reflect the tropopause height, and therefore, neither will the retrievals. Conversely, the other instruments are capable of measuring the height of the tropopause, more accurately defined by the temperature minimum. We avoided discussing the physics of the atmosphere in this section, because we believe the artifact is due more to the fall off of the TANSO-FTS sensitivity and that the vertical resolution of all instruments is not good enough to perfectly reflect the shape of the CH_4 VMR vertical profile at the tropopause. The tropopause heights of the instruments tend to agree within uncertainty, but when, for example, ACE-FTS only has observations every 3 or 5 km, that uncertainty is necessarily large. We are not willing to say that the tropopause heights are different because of this feature. We agree that differences in tropopause height pose significant challenges, but the most relevant argument for using pressure instead of tangent altitude is that the tangent altitudes for TANSO-FTS are not know and computing them introduces additional errors. We were careful to say “this feature indicates that the altitude at which this VMR decrease differs” and not that the tropopause heights differ.

44. p. 13 l. 24/28: By "below 90 hPa" do you mean less that 90 hPa or at lower altitudes? Similarly, does "Above 100 hPa" mean greater than 100 hPa, or at higher altitudes? Looking at the figure, the reader can deduce the appropriate answer but would benefit from less confusing wording.

Referring to pressure is correct in this context since we are discussing a figure with pressure on the axis. We have reworded the statements here to refer to levels.

45. p. 13 l. 24-27: This result seems to me as one of the most important in the manuscript and deserves elaboration. Does the variability have any notable features? Does it depend on sensitivity (s) or a priori influence? Did you find covariance with latitude (i.e. within the 30 degree bins), tropopause height, or season?

This is explored in Sect. 8. We specifically measure the relationship on mean VMR difference with latitude.

46. p. 13 l. 33-34/Figure 6: Do zonal differences exist in the unsmoothed data? How do the unsmoothed data fit into the goal mentioned in the introduction for assessing the applicability of Holl et al. (2016) to lower latitudes?

The unsmoothed data didn’t exhibit significant zonal differences, which is why we chose to show fewer figures. An explicit mention of the lack of zonal differences has been added to the text.

47. p. 13 l. 35-36: This sentence is confusing due to the the vague phrasing and verbosity (e.g. "actual differences one would expect"). Please reword.

The first half of this sentence has been re-written to read: “Fig. 6 shows the mean differences between the TANSO-FTS data product and those of other instruments.”

48. p. 13 l. 37: More consistent across instruments? Or across altitudes for each instrument?

The sentence is re-worded to explicitly name differences as those profiles shown in Fig. 6.

49. p. 14 l. 9-11: Why would the differences between ACE-FTS and ESA MIPAS be smaller? Does this shed light on differences between each of these instruments and TANSO?

The difference between ACE-FTS and ESA MIPAS is smaller, and expectedly so, because both instruments' data products are more accurate throughout the altitude range shown due their increased sensitivity and decade-long retrieval development.

50. Section 7.1: Does any of the methodology apply to Section 6 (e.g. criteria for minimizing the dominance of the prior) and vice versa?

No, only the selection of coincident measurements and smoothing is common. In Sect. 6, we compare measurements at each pressure level and use whatever pressure levels are available. In Sect. 7, we are integrating the vertical profiles and set a requirement of a minimum overlap of retrieval levels.

51. p. 14 l. 21-22/Figure 8: The way this is described, it seems contradictory with the caption on Fig. 8, "The vertical range of partial column integration varies for each pair of coincident profiles." If you mean that for each coincident measurement pair you match the vertical range of TANSO and each of the other vertical profiles, but that the vertical ranges across all coincident measurements vary, please describe this more explicitly somewhere in this section. Also, if that interpretation is correct, does the vertical range impact the distributions or correlations of the data? (This same question applies to Figure 9.)

There is some ambiguity here, that we have tried to correct. Each pair compared must have the same integration range. We didn't require every pair to have one range. That the magnitude of the partial columns varies with altitude is discussed on lines 19–24 and on page 15, lines 31–33, and statistics about the variability of altitude ranges are given in Table 3 and discussed on page 16.

52. p. 14 l. 34: Why a sensitivity threshold of 0.2? This seems a little low. The minimum of three pressure levels also seems low unless they are contiguous (i.e. don't skip filtered out data in the profile). If the data points do not adjoin each other, did you apply criteria on how far apart the levels can be?

This is partly because the sensitivity of TANSO-FTS is this low. The problem isn't just the sample size, however, with a higher threshold, the pairs that remain in the comparison tend to to have short vertical ranges of integration. We are comparing smoothed data which are necessarily interpolated, so the altitudes of overlap are contiguous.

53. p. 14 l. 5-7: How different are the results when these 23for the outliers in Figures 8 and 9?

There are no outliers, if they are shown, they lie beneath the "good" data. They all have small partial column values because the integration interval is small. This comment is a little confusing because, since we selected a sensitivity threshold, the 23% of the data that does not meet the criteria are excluded by the criteria. A comment about the behaviour of the excluded data has been added to the text.

54. p. 15 l. 10-20: Given that the values are all on a pressure grid, what is the advantage of integrating in altitude/z versus pressure/P? Do you account for water vapour (i.e. use "dry" P/T)?

Reformulating the problem as an integral of pressure introduces a dependence on the vertical profile of the mass density ($dz = dP/(-\rho g)$). We did not convert to “dry” P–T because the coincident measurements are co-located so the humidity should be close to the same, but also, introducing a measurement of water vapour introduces the error in that measurement.

55. p. 15 l. 27: If partial columns with large gaps in the vertical are included (ref. my comment on p. 14 l.34), an uncertainty related with the interpolation should also be propagated through the calculation.

The data are smoothed and necessarily interpolated. The uncertainty of the interpolated data accounts for the interpolation process.

56. p. 15 l. 29/Figures 8 and 9: Given the emphasis of the paper on zonal dependencies, please add a colour scale to each plot associated with latitude (bin)?

Unfortunately, colourizing Figs. 8 and 9 with latitudes would not be helpful for our paper. You wouldn’t see a latitudinal dependence on the correlation, but, you still see clustering of the data. Since the partial column depends on integration range and the lower limits of the ACE-FTS and MIPAS observations depend on latitude (water vapour, opacity and clouds increasing towards the tropics), then the latitudinal dependence in the partial columns is due to the observation geometry of those two satellites.

57. p. 15 l. 35-37: What would cause a bias in the intercept? Altitude range? Spectroscopy?

A non-zero, positive y-intercept means that the y-axis data, TANSO-FTS partial columns, are generally larger, by this amount, than those from other instruments. This is consistent with Fig. 5 which shows that the TANSO-FTS VMRs are generally greater than the other data products.

58. p. 16 l. 8/Table 3: Why is the minimum altitude for the NDACC measurements so high? Figure 3 indicates that TANSO is at least somewhat sensitive closer to the surface that 3km.

3.3 km is actually very low for a remote sensing instrument. This value not for the NDACC stations, but for the overlap between TANSO-FTS and NDACC obeying our sensitivity and flagging criteria. The pressure maximum in the TANSO-FTS retrievals is always less than that of NDACC, and the sensitivity of TANSO-FTS falls off at the lower altitudes, as shown in Fig. 3e.

59. p. 16 l. 26: When combining the results, are all data weighed equally, or do you take into account the uncertainties of measurements? Is this the average across all latitudes, or is it a bias that is consistent for all latitudes? Have you also assessed whether altitude related biases exist in the combined data?

The least-squares regression is weighted, since each measurement has a unique uncertainty estimate. Each instrument is treated equally, so that with the largest sample size (MIPAS) has the highest contribution. Notes about the regression have been added to the text. Altitude-related biases exist due to the averaging kernels of TANSO-FTS. These can be inferred from Fig. 5.

60. p. 16 l. 28-29: Did you find a bias in the sub-tropics or mid-latitudes?

The bias is latitude dependent; it is the result of weighted least squares regression of all the data, which returned a slope and intercept that are statistically significant. The values discussed are approximations based on the regression results. The magnitude of the bias varies between the tropics and the poles. It is not clear enough in the manuscript that these results are from the least-squares regression and we have attempted to correct that.

61. p. 16 l. 31-32: It is not clear what type of comparison was done? Regression? ANCOVA?

We compared each parameter with our results, as in Fig. 10, and looked at least squares regression results and correlation statistics. This has been specified in the test.

62. p. 17 l. 13: The mismatch in vertical extent you point out seems to indicate that these other satellites are not appropriate/useful for validation. If this is not your argument, please rephrase this sentence to make the argument more clear.

The vertical mismatch is unfortunate, and MIPAS and ACE-FTS are not the perfect data sets to compare to, but finding a comparable, satellite-based instrument with the same data product and an overlapping operation timeline is not so easy. This problem is also a result of our TANSO-FTS analysis and comparison. The sixteen ground-based observatories do have fully overlapping altitude ranges. We agree that this sentence and the preceding one are poorly written and do not convey our intent, so they have been rephrased.

63. p. 17 l. 15: Have you tried smoothing the TANSO profiles to NDACC to see if the agreement is robust?

Because NDACC has a finer vertical resolution than TANSO-FTS TIR, and is sensitive to more vertical structure in the profiles, this operation may not result in “smoothing.” Furthermore, since the TANSO-FTS sensitivity is low and its retrieved profiles are close to the a priori over most of its altitude range, the result would be mostly mixing the TANSO-FTS a priori and the NDACC a priori together. Holl et al., (2016) measured the vertical resolution for each pair of measurements to determine which was the lower resolution instrument and in a small number of cases smoothed the TANSO-FTS profiles with the NDACC averaging kernels and a priori. We felt it was more important to maintain a consistent set of steps to create the data sets presented in Figs. 5, 6, 8, 9 and 10.

64. p. 17 l. 18-19: Given these biases, would you recommend "calibrating" the TANSO retrievals?

This would be case dependant, and because of the small magnitude of the bias, up to the researcher’s discretion. Due to the spread of the data shown in Fig. 10 and the uncertainty of each instrument’s retrieval, we would not recommend calibrating individual profiles, but would recommend that the bias is considered when a large number of profiles are combined.

65. p. 17 l. 19-21: Please include how this altitude feature varies (or doesn't) with altitude.

We are not certain of what the reviewer is asking us here. The feature is at a specific altitude, it does not vary with altitude, it is not present at other altitudes. We have re-written this sentence to ensure that which profiles are discussed is clear, and that we are discussing Fig. 5 is also clear.

66. p. 17 l. 25: It is not clear what "taken over altitude and latitude" means; please reword.

This refers to Fig. 10, but is a typo and should read "dependence of the mean differences, taken over altitude, on latitude." This is cumbersome and has been revised.

67. p. 17 l. 26-27: What improvements are expected in future versions of the retrieval (e.g. priors, spectroscopy)? Based on your results, what would you recommend needs the most/least attention to produce a more accurate data product? Given the limitations of this TANSO product, what applications would it be suited to?

We have asked our co-authors who work on the TANSO retrievals what is next and included some comments in the conclusion about this. The data product is very important because of the altitude range that it covers, which covers the middle and upper troposphere – not covered by the other data products. This differs and compliments ground based observatories in it global spatial coverage. This is mentioned earlier in the conclusion, but maybe not strongly enough. We have added a statement at the beginning of the conclusion pointing out the importance of the data product.

2.4 Technical suggestions

Comments referring to the addition or removal of punctuation were included where I thought they might improve readability and are thus suggestions rather than corrections, except in cases where the serial comma in a list needs to be added.

1. p. 1 l. 3: Change "CO2 and CH4" to "carbon dioxide (CO2) and methane (CH4)"

We added definitions for $\mathrm{CH_4}$ and $\mathrm{CO_2}$.

2. p. 1 l. 15: Change "examine" to "examining"

The subject in this list is "we," so examine is correct here. We have added an "and" to better join clauses.

3. p. 2 l. 3-6: Add VMR (parts per notation) in parentheses after listed percentages.

Added $\pm 40\,\mathrm{ppbv}$ in parentheses.

4. p. 2 l. 5: Change "investigated" to "investigate"

Made the correction, thank you.

5. p. 2 l. 7: Unless "over the equator" is specifically what is meant, change to "in the tropics"

Made the change.

6. p. 2 l. 29: Remove the comma after "local"

Made the change.

7. p. 2 l. 33: Change "made in coincidence" to "coincident"

Made the change.

8. p. 3 l. 30: Add a comma after "surface temperature"

Made the change.

9. p. 3 l. 32: Add a comma after "(Maksyutov et al., 2008; Saeki et al., 2013)"

Made the change.

10. p. 4 l. 9: Remove dash after "ACE"

This is the ACE-MAESTRO instrument.

11. p. 4 l. 19: Add comma after "Boone et al. (2005)"

Made the change.

12. p. 4 l. 24: Remove comma after (Picone et al., 2002)"

Made the change.

13. p. 4 l. 25: The use of "and" does not seem correct. Could be replaced with "assimilated into" or "from" depending on the relationship between the met data and the model.

Changed “and” to “‘with.”

14. p. 4 l. 30: Add comma after "profiles"

Made the change.

15. p. 4 l. 35: Add comma after "Odin"

Left as is.

16. p. 5 l. 7: Awkward placement of "(inclination of 98)"

This has been removed.

17. p. 5 l. 9: Add comma after "cloud parameters"

Made the change.

18. p. 5 l. 11: Add comma after "2004"

Made the change.

19. p. 5 l. 12: Reword end of this sentence, e.g. change ", but" to "with"

Changed to “but with.”

20. p. 5 l. 24: Add comma after "limb scan"

Made the change.

21. p. 5 l. 26: Add comma after "temperature"

Made the change.

22. p. 6 l. 20: Change the commas around "below 25km" to parentheses.

Made the change.

23. p. 7 l. 7: Rephrase "dynamical nature" to a more precise term.

Changed it to “characteristic atmospheric dynamics over Antarctica.”

24. p. 7 l. 15-16: The use of and phrasing after the semi-colon is awkward and makes the sentence unclear.

This is the correct use of a semi-colon, separating two list items which contain commas within them.

25. p. 8 l. 11-12: The parenthetical is awkwardly worded; please revise for clarity.

This sentence has been revised for clarity.

26. l.12: Change "differences is also" to "differences are also"

Correction made.

27. p. 8 l. 13: Please reword "When examining dates with several measurements" to make more clear.

Revised to “several measurements made on the same day.”

28. p. 8 l. 16-17: The grammatical structure of this sentence is difficult to follow. Please reword.

Revised this sentence.

29. p. 9 l. 32: Add comma after "differ"

Made the change.

30. p. 10 l. 10: Change comma after "kernel" to semi-colon.

Replaced with a period.

31. p. 10 l. 13: Add comma after "role"

Made the change.

32. p. 10 l. 20: Add comma after "altitudes"

Left as is.

33. p. 10 l. 22: Change "70km. This is shown in Fig. 3e." to "70km (Fig. 3e)."

Moved this clause earlier in the text.

34. p. 10 l. 22: Add comma after "development"

Made the change.

35. p. 10 l. 27-28: Please reword. The structure of this sentence is difficult to follow, e.g. the verb ("are shown") appears twice and each has its own modifier.

This sentence had remnants of earlier changes, typos fixed, and further revisions made, thank you.

36. p. 11 l. 22-23: Remove parentheses, and add a period or semi-colon after "retrieval"

Made the change, using a period.

37. p. 11 l. 24: Please state which retrieval you're referring to (seems like the higher-resolution profile, but not self-evident).

Revised this clause.

38. p. 11 l. 27: Remove comma after "equilibrium"

Left as is.

39. p. 12 l. 21-22: Change "looking for" to "filtering" and remove ", and then filtering these events," to make the sentence more clear.

Made the change.

40. p. 13 l. 20: I think a verb is missing after "VMR decrease" (e.g. "occurs").

Yes, there was a missing word.

41. p. 13 l. 33: Add a comma after "zonally"

Made the change.

42. p. 16 l. 29: Change "or" to "and" ("0.014 ppmv or 0.020 ppmv"), and "Pole" should be plural.

Made the change.

43. p. 17 l. 10: Remove commas after "sensitivity" and "product"

Made the change.

44. p. 17 l. 12: Remove comma after "altitudes"

Left as is.

45. p. 17 l. 12-13: Please reword "and that there is a limitation on the useful upper altitude of its data product of below 15 or 20km" to follow the clarity and structure of the beginning of the sentence.

This has been revised.

46. p. 17 l. 14: Add "upper" before "troposphere" (Without the addition, this sentence is misleading.)

Made the change.

47. p. 17 l. 20: Phrase starting with "and in a consistent manner" needs rewording for clarity.

This has been revised.

48. Figures 2 and 10: The legend has two icons for every instrument, which adds extra visual clutter.

Left as is.

49. Figure 3: Several line colours do not appear in the legends of a-d. Instead of using a legend, you might consider labelling each line with the pressure using the same colour for the text.

There is not enough room to label every profile, so we only labelled 5 kernels for each instrument.

50. Figure 7: A heat map or similarly sequential colour scheme could be more helpful for this type of plot.

The bin widths are important here, smoothing the data would not really be helpful.

51. Figure 8: The “R” is missing on the R2 line of each sub-figure, and it looks like some other letters and numbers might also be missing.

These, and other characters in other figures were stripped from the pdf files when the AMT header was added. We will have to work with editors to ensure these characters remain in a revised manuscript.

52. Figure 10: The degree symbol is missing between parentheses on the x-axis label. Also, please add additional tick marks on the x-axis. You might consider including light gray grid lines behind the data.

Left background as is, will talk to editors about missing symbols.

3 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

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)

Abstract. The primary instrument on the Greenhouse gases Observing SATellite (GOSAT) is the Thermal And Near infrared Sensor for carbon Observations (TANSO) Fourier Transform Spectrometer (FTS). TANSO-FTS uses three short-wave infrared (SWIR) bands to retrieve total columns of CO_2 and CH_4 along its optical line-of-sight, and one thermal infrared (TIR) channel to retrieve vertical profiles of CO_2 and CH_4 volume mixing ratios (VMRs) in the troposphere. We examine version 1 of the TANSO-FTS TIR CH_4 product by comparing co-located CH_4 VMR vertical profiles from two other remote sensing FTS systems: the Canadian Space Agency's Atmospheric Chemistry Experiment-FTS (ACE-FTS) on SCISAT (version 3.5), and the European Space Agency's Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) on Envisat (ESA ML2PP version 6 and IMK-IAA reduced-resolution version V5R_CH4_224/225), as well as 16 ground stations with the Network for the Detection of Atmospheric Composition Change (NDACC). This work follows an initial inter-comparison study over the Arctic, which incorporated a ground-based FTS at the Polar Environment Atmospheric Research Laboratory (PEARL) at Eureka, Canada, and focuses on tropospheric and lower-stratospheric measurements made at middle and tropical latitudes between 2009 to 2013 (mid 2012 for MIPAS). For comparison, vertical profiles from all instruments are interpolated onto a common pressure grid, and ~~the~~ smoothing is applied to ACE-FTS, MIPAS, and NDACC vertical profiles~~are smoothed using the~~. 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 ~~averaging kernels~~a priori and averaging kernels

in all cases. We present zonally-averaged mean CH_4 differences between each instrument and TANSO-FTS with and without smoothing, and examine their information content, sensitive altitude range, correlation, a priori dependence, and the variability within each data set. Partial columns are calculated from the VMR vertical profiles, and their correlations are examined. We find that the TANSO-FTS vertical profiles agree with the ACE-FTS and both MIPAS retrievals' vertical profiles within 4 % ($\pm\sim 40$ ppbv) below 15 km when smoothing is applied to the profiles from instruments with finer vertical resolution, but that the relative differences can increase to on the order of 25 % when no smoothing is applied. Computed partial columns are tightly correlated for each pair of data sets. We investigated investigate whether the difference between TANSO-FTS and other CH_4 VMR data products varies with latitude. Our study reveals a small dependence of around 0.1 % per ten degrees latitude, with smaller differences over the equatortropics, and greater differences towards the poles.

1 Introduction

The Greenhouse gases Observing SATellite (GOSAT) was developed by Japan's Ministry of the Environment (MOE), National Institute for Environmental Studies (NIES), and the Japan Aerospace Exploration Agency (JAXA), and was launched in 2009 with an inclination of 98° (Yokota et al., 2009). The objectives of the GOSAT mission include monitoring the global distribution of greenhouse gases, estimating carbon dioxide (CO_2) source and sink locations and strengths, and verifying the reduction of greenhouse gas emissions, as mandated by the Kyoto Protocol. GOSAT carries two instruments: the Thermal And Near infrared Sensor for carbon Observations (TANSO) Fourier Transform Spectrometer (FTS) and the TANSO Cloud and Aerosol Imager (TANSO-CAI). 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.

TANSO-CAI is a radiometer with four spectral bands between 0.37 and 1.6, each around 0.02wide and chosen to avoid and absorption. TANSO-CAI that is able to measure the cloud fraction in the field-of-view of TANSO-FTS (Ishida and Nakajima, 2009; Ishida et al., 2011). TANSO-FTS is a nadir-viewing double-pendulum FTS, whose technical details are described in Sect. 2.1. TANSO-FTS makes observations of infrared radiation emitted from the Earth's atmosphere in four bands. Three bands are in the short-wave infrared region and are used to measure total columns of CO_2 and methane (CH_4). The fourth channel is in the thermal infrared (TIR) to provide GOSAT with sensitivity to the vertical structure of CO_2 and CH_4.

This work follows Holl et al. (2016), who compared Atmospheric Chemistry Experiment (ACE) FTS version 3.5 (v3.5) and TANSO-FTS TIR version 1 (v1) vertical profiles with those measured by a ground-based FTS at the Polar Environment Atmospheric Research Laboratory (PEARL) at 80° N in Eureka, Canada (Batchelor et al., 2009) . We employ a similar methodology, extend that study globally, and include multiple ground-based FTSs that are part of the Network for the Detection of Atmospheric Composition Change (NDACC) (Kurylo and Zander, 2000) . Holl et al. (2016) observed that after smoothing the ACE-FTS profiles using the TANSO-FTS averaging kernels and a priori profiles, the difference is close to zero above 15 km, but that there is a bias at lower altitudes where TANSO-FTS retrieves more CH_4, with a mean excess of 20 ppbv in the troposphere.

The data analyzed by Holl et al. (2016) are limited to a single location characterized by cooler temperatures and lower humidity than lower latitudes, and limited latitudinal transport. Our objective is to investigate whether the results of Holl et al. (2016) are local , or hold at all latitudes, and to provide additional global validation of the TANSO-FTS v1 CH_4 data product. Any biases in the v1 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.

In this manuscript, we examine the TIR data product from TANSO-FTS, specifically, CH_4 volume mixing ratio (VMR) vertical profiles, by determining when TANSO-FTS TIR retrievals of CH_4 were made in coincidence with those of other satellite-borne and ground-based FTS instruments. Comparisons of satellite instruments are made with the ACE-FTS on SCISAT, described in Sect. 2.2, and the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) on the Environmental Satellite (Envisat), described in Sect. 2.3. The NDACC InfraRed Working Group (IRWG) has a network of ground-based FTSs; we used 16 that retrieve vertical profiles of CH_4 VMR to compare with the TANSO-FTS TIR data. The NDACC data are described in Sect. 2.4. A summary of the instruments used in this study is given in Table 1.

The question we are asking in this validation study is not *what is the magnitude of the difference between retrieved* CH_4 *vertical profiles from TANSO-FTS and other instruments*, 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.

For each comparison pair, the averaging kernels, information content, and variability of the retrievals are examined in Sects. 3 and 5. The instrument with finer vertical resolution is smoothed using the averaging kernels of the instrument with coarser vertical resolution (TANSO-FTS in all cases presented here) in order to account for the structure intrinsic to a finer-resolution instrument. For each coincident pair, the absolute and relative differences of the smoothed and unsmoothed VMR vertical profiles are found and their means, correlation coefficients, R^2, and numbers of coincident pairs are computed at each pressure level. For each vertical profile in a coincident pair, an overlapping vertical extent is selected using the sensitivity, or response, of the TANSO-FTS retrieval (area of the averaging kernel matrix), partial columns are computed over this range, and their correlations are examined. Finally, this altitude range is used to estimate the mean VMR difference taken over the vertical range for each coincident pair of profiles. This dataset data set shows any biases related to latitude, or any other parameters of the TANSO-FTS retrieval, such as incidence angle or surface type (land or water).

Sect. 4 describes the methods and criteria for determining coincident measurements between TANSO-FTS and each instrument. Sect. 6.1 provides a detailed description of the comparison methodology. Comparison results for each instrument are presented in Sect. 6.2. The satellite instruments are zonally averaged and each NDACC site is shown. Partial column calculation methodology is presented in Sect. 7.1 and correlation results are shown in Sect. 7.2. A discussion follows in Sect. 8, focusing on our investigation of biases within the TANSO-FTS retrievals related to latitude and other parameters.

2 Data sets

2.1 TANSO-FTS

TANSO-FTS makes measurements of radiance in four bands~~: 12900–13200, 5800–6400, 4800–5200, and~~ ; the TIR band is between 700–1800 cm^{-1} ~~. The fourth band is in the TIR~~ and is used to retrieve vertical profiles of CH_4 VMRs. TANSO-FTS has a spectral resolution of 0.2 cm^{-1} and operates in a nadir or near-nadir viewing geometry (Kuze et al., 2009). To improve coverage, its field of view sweeps longitudinally, and TANSO-FTS makes several measurements along each cross track, five measurements prior to August 2010, and three since then (Kuze et al., 2012). This leads to TANSO-FTS having the highest density of measurements and greatest spatial coverage among the instruments considered herein.

Retrievals of v1 CH_4 follow the ~~methodology~~ nonlinear maximum a posteriori method used for v1 CO_2 presented in Saitoh et al. (2009, 2016). They are performed on a fixed pressure grid and the pressure levels are adjusted based on the averaging kernels for the retrieval. In the v1 retrieval algorithm, water vapour, nitrous oxide, ozone concentrations, temperature, surface temperature, and surface emissivity were retrieved simultaneously with CH_4 concentration from V161.160 L1B spectra. A priori data are based on simulated data from the NIES transport model ~~(Maksyutov et al., 2008; Saeki et al., 2013)~~ (TM) (Maksyutov et al., 2008; Saeki et al., 2013) , and the retrievals use the HITRAN 2008 ~~linelist~~ line list (Rothman et al., 2009) with several updates up to ~~2011.~~ 2011 (Saitoh et al., 2009) .

An initial comparison of TANSO-FTS v1 to a single NDACC station, Eureka, and to ACE-FTS measurements made in the Arctic within a quadrangle surrounding PEARL (60–90° N and 120–40° W) has been recently made (Holl et al., 2016). The v1 CH_4 product was also compared globally with the version 6 CH_4 data product from the Atmospheric Infrared Sounder (AIRS) on Aqua (Zou et al., 2016).

2.2 ACE-FTS

ACE-FTS was launched into ~~a circular~~ low-Earth orbit ~~with an inclination of 74~~in 2003 ~~onboard~~ on-board the Canadian Space Agency's (CSA's) SCISAT. ~~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 285–1030and a spectral resolution of 1–2.~~ The scientific objectives of ACE are to study ozone distribution in the stratosphere, the relationship between atmospheric chemistry and climate change, the effects of biomass burning on the troposphere, and the effects of aerosols on the global energy budget (Bernath, 2017).

ACE-FTS is a high-resolution, double-pendulum FTS with a spectral resolution of 0.02 cm^{-1} that covers a broad spectral range between 750–4400 cm^{-1}. It operates in solar occultation mode, making a series of measurements for tangent altitudes down to 5 km (or cloud tops) at local sunrise and sunset along its orbital path (Bernath et al., 2005). Its level 2 data products are vertical profiles of temperature, pressure, and the VMRs of 36 trace gases, as well as isotopologues of major species, reported on an altitude grid at the measurement tangent altitudes or interpolated onto a 1 km grid. Retrievals of the Version 2.2 (v2.2) data product are described in Boone et al. (2005), and updates regarding the latest release, Version 3.5 (v3.5), are described in Boone et al. (2013). V3.5 retrievals, with the data quality flags (v1.1) described in Sheese et al. (2015), are used herein.

When performing trace gas retrievals, tangent altitudes for each observation and vertical profiles of temperature and pressure are also retrieved using spectral fitting (not simultaneously). Comparisons with TANSO-FTS are made on a pressure grid using the retrieved pressure values at the ACE-FTS measurement heights. A priori temperature and pressure for ACE-FTS are derived from the NRL-MSISE-00 model (MSIS) (Picone et al., 2002) , and from meteorological data provided by the Canadian Meteorological Centre and with their Global Environmental Multiscale (GEM) model (Côté et al., 1998). Fitted spectra are computed using the HITRAN 2004 spectral linelist line list (Rothman et al., 2005) with modifications described in Boone et al. (2013).

Validation of v2.2 CH_4 VMR vertical profiles is presented in de Mazière et al. (2008) and was performed using several ground-based FTSs that are part of NDACC, as well as one at Poker Flat. For that comparison, partial columns were computed from the ACE-FTS CH_4 profiles, and the correlation between partial columns computed from ground-based FTSs and from ACE-FTS was investigated. Validation was also done against the balloon-borne SPIRALE (Spectroscopie Infra-Rouge d'Absorption par Lasers Embarqués), the Halogen Occultation Experiment (HALOE) on the Upper Atmosphere Research Satellite, and MIPAS. de Mazière et al. (2008) determined that the ACE-FTS v2.2 CH_4 data are accurate to within 10 % in the upper troposphere and lower stratosphere and to within 25 % at high altitudes. More recently, Jin et al. (2009) compared CH_4 from the Canadian Middle Atmosphere Model (CMAM) with measurements from ACE-FTS, the Sub-Millimeter Radiometer (SMR) on Odin and the Microwave Limb Sounder (MLS) on Aura, and found agreement with ACE-FTS within 30 %. Updates to the ACE-FTS validation effort using v3.0 data and a description of the differences between v2.2 and v3.0 are presented in Waymark et al. (2013). Waymark et al. (2013) found a slight reduction in CH_4 VMR in the v3.0 data near 23 km, and a larger reduction of around 10% between 35–40 km.

2.3 MIPAS

MIPAS is a limb-sounding FTS that was placed in polar (inclination of 98) low-Earth orbit in 2002 onboard the European Space Agency's (ESA's) Envisat. MIPAS aimed to provide global observations, during both night and day, of changes in the spatial and temporal distributions of long- and short-lived species, temperature, cloud parameters, and radiance. The instrument was intended to have a maximum spectral resolution of 0.025 cm^{-1} (Fischer et al., 2008), but the slide system for the interferometer mirrors encountered a problem in 2004, and observations used in this study were made with a reduced effective spectral resolution of 0.0625 cm^{-1}, but with finer vertical sampling. Further complications arose in 2012 and ESA lost communication with Envisat, ending the mission.

The spectral range of MIPAS is 685–2410 cm^{-1}, allowing the retrieval of multiple trace gases. MIPAS spectra are processed independently by four research groups (Raspollini et al., 2014). In this paper, we consider two: the ESA operational analysis and the Karlsruhe Institute of Technology Institute of Meteorology and Climate Research (IMK) and the Instituto de Astrofísica de Andalucía (IAA) analysis, both described in the following subsections.

2.3.1 ESA MIPAS

We use MIPAS Level 2 Prototype Processor version 6 (ML2PP v6) of the ESA operational analysis. Early versions of the ESA MIPAS gas retrievals are described in Raspollini et al. (2006) (full-resolution Instrument Processing Facility version 4.61 (IPF v4.61)) and the ML2PP v6 upgrades and reduced resolution adaptations are described in Raspollini et al. (2013). Retrievals are made using a global fitting scheme followed by a posteriori Tikhonov regularization with self-adapting constraints (Raspollini et al., 2013). The ML2PP v6 data provide retrieved VMR vertical profiles of ten atmospheric gases between approximately 6 to 70 km. Temperature and pressure are retrieved from the spectra at each tangent point of a limb scan and a corresponding altitude grid is built from the lowest engineering tangent altitude using the equation of hydrostatic equilibrium. Initial guesses for vertical profiles of a target trace gas, temperature and interfering species are the weighted average of the results from the previous scan, an appropriate merging of IG2 (initial guess 2) climatological profiles (Remedios et al., 2007), and, if available, data from the European Centre for Medium-range Weather Forecasts (ECMWF). Spectra are computed using a specialized ~~linelist~~ line list derived from HITRAN 1996 (Rothman et al., 1998).

The IPF v4.61 CH_4 data product has been validated by Payan et al. (2009) against four balloon instruments, including SPIRALE, three aircraft instruments, six ground-based FTSs (all are considered herein), and HALOE. They found good agreement with a 5 % positive bias in the lower stratosphere and upper troposphere. ML2PP v6 CH_4 was compared with BONBON air sampling measurements by Engel et al. (2016). The reduced-resolution CH_4 measurements (2005–2012) agree with in situ data within 5–10 %. CH_4 (and N_2O) from ESA MIPAS have been assimilated by the BASCOE code and the assimilated products have been compared with MLS and ACE-FTS (Errera et al., 2016). The analysis has proven the high quality of the MIPAS data, but it has also identified the presence of some outliers, especially in the tropical lower stratosphere, and some discontinuities due to issues in the measurements.

2.3.2 IMK-IAA MIPAS

The IMK-IAA MIPAS retrieval algorithm has been developed to include and account for deviations from local thermal equilibrium. The data presented here are IMK-IAA reduced-resolution version V5R_CH4_224/225. The early retrieval algorithms are described by von Clarmann et al. (2009), and the updates made to the current version are described by Plieninger et al. (2015). Temperature and tangent altitude are retrieved from the spectra, and pressure is computed from the equation of hydrostatic equilibrium. V5R_CH4_224/225 uses the HITRAN 2008 ~~linelist~~ line list (Rothman et al., 2009). Temperature a priori profiles are determined from ECMWF analyses and MIPAS engineering information. The IMK-IAA retrieval uses Tikhonov first-order regularization in combination with an all-zero CH_4 a priori profile, which serves to smooth the profiles.

Validation of the IMK-IAA MIPAS V5R_CH4_222/223 data has been presented in Laeng et al. (2015). They compare data against ACE-FTS, HALOE, the MkIV balloon FTS, the Solar Occultation For Ice Experiment (SOFIE) on the Aeronomy of Ice in the Mesosphere (AIM) satellite, the SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) on Envisat, and a cryogenic whole-air sampler (collects gas bottle samples during aircraft flights). They found an agreement within 3 % in the upper stratosphere with other satellite instruments, but in the lower stratosphere ~~,~~ (below 25 km~~,~~

) a high bias was found in the MIPAS retrievals of up to 14 %. The V5R_CH4_224/225 has more recently been validated by Plieninger et al. (2016), using ACE-FTS, HALOE, and SCIAMACHY. They found MIPAS CH_4 retrievals to be larger by around 0.1 ppmv below 25 km, or around 5 %.

2.4 NDACC

NDACC is a global network of a variety of instruments that provides measurements of tropospheric and stratospheric gases that are directly self-comparable (Kurylo and Zander, 2000). The network consists of over 70 stations sparsely distributed at all latitudes. Information about NDACC is available at www.ndacc.org. In this work, we only consider a small subset of NDACC stations that feature high-resolution FTSs and provide a CH_4 VMR vertical profile data product via the NDACC data base. Sepúlveda et al. (2012, 2014) demonstrated the good quality of CH_4 profiles that can retrieved from the NDACC FTS measurements. The stations are listed in Table 1, along with their locations, ~~references, and some information about each instrument~~spectral range and resolution, and references.

The stations do not use identical instruments, spectroscopic lines, or retrieval methods. All but one station use a version of a Bruker 120/5 M or HR, and have predominantly adopted, or upgraded to, the Bruker 125HR. Some stations have more than one instrument, and the type of instrument has changed over time at many of the stations. Toronto, 43.6° N, uses a Bomem DA8.

Retrievals are generally performed using either PROFFIT (Hase et al., 2004) or SFIT4 (Pougatchev et al., 1995) following harmonized retrieval settings recommended by the NDACC IRWG (Sussmann et al., 2011, 2013). Data used herein are from the NDACC database. A summary of retrieval settings is provided by Bader et al. (2017). Lauder and Arrival Heights, at 45.0° S and 77.8° S, use a retrieval strategy that adheres to that defined in Sussmann et al. (2011), with a relaxed Tikhonov regularization constraint at Arrival Heights due to the ~~dynamical nature of the Antarctic atmosphere~~characteristic atmospheric dynamics over Antarctica. Jungfraujoch, at 46.6° N, uses SFIT2. It has been established within the NDACC IRWG that the regularization strength of the CH_4 retrieval strategy should be optimized so that the number of degrees of freedom for signal (DOFs) is limited to approximately 2 (Sussmann et al., 2011).

3 Data set variability

To provide context for the VMR differences found when comparing each instrument to TANSO-FTS, shown in Sect. 6, we have examined the ~~internal variability~~ 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. ~~Such differences include the~~ For example, TANSO-FTS and MIPAS have a much higher data density ~~of TANSO-FTS and MIPAS compared to~~ than ACE-FTS, which only makes two sets of observations per orbit~~; and that the NDACC stations are stationary~~.

Following Holl et al. (2016), we are aware that TANSO-FTS CH_4 retrievals are dependent on the a priori used, especially at high altitudes. ~~To examine the variability within the~~ TANSO-FTS ~~data set, we examined the distribution of VMR vertical profiles and of the difference between the retrieval and~~ 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~~, to determine whether the variability of the TANSO-FTS retrievals matched that of their~~. This is indicative of the instrument sensitivity discussed in Sect. 5 and shows by how much the retrievals deviate from the a priori. We examined 3000 randomly selected TANSO-FTS measurements by interpolating the a priori and retrieved profiles to the pressure grid used in our comparisons (Sect. 6.1), then computed the difference between the retrieval and the a priori at each pressure level, and their mean and standard deviation. Fig. 1 shows the mean ± 1 standard deviation of the difference between the TANSO-FTS CH_4 retrievals and their corresponding a priori profiles. The peak value is ~~0.03~~30 ppbv near 10 km ($\sim$1.5 %) with a standard deviation of the same magnitude.

To examine the variability of the ACE-FTS CH_4 data product, we compared each retrieved profile from an ACE-FTS sunset/sunrise ~~measurement in a year to the sunset/sunrise measurement acquired on the~~ (occultation direction) to that from the next orbit, taking care to avoid a comparison between sunset and sunrise occultations (which are in different hemispheres), or when an acquisition was not ~~made during an~~ recorded during a subsequent orbit. Considering all sunset occultations in 2011, there were 1402 retrieved vertical profiles, and 820 sequential pairs. These pairs are separated by 97 minutes and have a mean spatial separation of 1180 ± 20 km, depending on the latitude of the measurement. For each pair, we computed the ~~difference between pairs of VMRs~~ VMR difference on the ACE-FTS 1 km tangent altitude grid, and then found the mean and standard deviation, which are shown in Fig. 1. Within the ACE-FTS data, the largest systematic variability (~~−0.004~~−4 ppbv) occurs around 30 km, with extreme outliers being observed at the lowest tangent altitudes. The mean magnitude of the ACE-FTS variability is ~~0.002~~2 ppbv (0.1 %) at all altitudes, and ~~0.009~~9 ppbv below 15 km (0.4 %).

To examine the variability of the MIPAS data sets, we compared the vertical profiles retrieved by IMK-IAA and ESA that ~~are~~ were made from the same ~~set of~~ MIPAS limb observations and within our coincident data set. This provides an indication of the impact of different retrieval algorithms on retrieved profiles. For each pair of retrieved vertical profiles from a single set of MIPAS spectra, we interpolated the ESA retrieval to the IMK-IAA 1 km grid and computed their difference (IMK-IAA − ESA), and then found the mean and standard deviation. Fig. 1 shows the mean ± 1 standard deviation for this comparison. The two retrievals show good agreement above 30 km (not shown), while the IMK-IAA data has a positive bias relative to the ESA data product of around 0.15 ppmv between 20 and 30 km. This bias is consistent with the validation results presented in Laeng et al. (2015). The ESA and IMK-IAA comparison exhibits the largest variability, with a mean magnitude (mean of absolute values) of ~~0.05~~50 ppbv (2 %) for the altitude range considered (9–34 km). Since the two products use the same spectra, it is possible that part of the internal instrument variability is hidden in this approach.

To investigate the variability of the NDACC data, we ~~considered the~~ compared pairs of observations made at an NDACC site on the same day. We considered only NDACC CH_4 VMR vertical profiles that were in coincidence with TANSO-FTS~~and~~

~~found a subset~~. For each pair of NDACC measurements~~for each site that were made on the same day. We then~~, we computed the CH_4 VMR differences ~~for each pair of measurements made on the same day~~ on the standard NDACC retrieval grid (earlier profile minus later; if there are multiple ~~profiles, the differences are all~~ coincidences in a day, differences are found relative to the earliest). The mean and standard deviation of these differences ~~is~~ are also shown in Fig. 1. When examining ~~dates with several measurements~~ several measurements from the same day, the NDACC differences show a systematic mean increase in tropospheric CH_4 with time during a single day. This variability is small, however, with a mean of ~~-0.004~~ -4 ppbv below 30 km and a peak at 12 km of ~~-0.006~~ -6 ppbv (0.3 %).

~~Of the four datasets, our~~ Our variability investigation found that the ACE-FTS data ~~exhibits~~ exhibit the smallest variability between measurements, ~~and~~ that MIPAS exhibits the largest, ~~while~~ and that NDACC and TANSO-FTS are of similar magnitudes. The magnitude of the internal variability of the data sets is between ~~0.05~~ ± 2 ppbv (e.g., for NDACC and ACE-FTS in the upper troposphere) and ~~0.004~~ ± 3 ppbv, or around 2 % (e.g., for TANSO-FTS ~~in the upper troposphere and when comparing ESA and IMK-IAA MIPAS~~and the lower limits of ACE-FTS).

4 Coincidences

Due the coverage and data collection rates of each instrument, different coincidence criteria were used. ~~In the case of~~ ACE-FTS ~~, which only records~~ 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, ~~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. For ACE-FTS and NDACC, we sought measurements made within 12 hours and within 500 km of each TANSO-FTS measurement (spatial separation calculated using the Vincenty method (Vincenty, 1975)). For the MIPAS data sets, we sought measurements made within 3 hours and 300 km. When searching for MIPAS–TANSO-FTS coincidences within 12 hours and 500 km, we find approximately 180,000 coincidences per month.

The criteria used in this study are comparable to previous CH_4 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.

TANSO-FTS CH_4 VMR vertical profiles tend not to be sensitive above the upper troposphere (see Sect. 5), while ACE-FTS and MIPAS retrievals have a limited vertical extent in the troposphere. To ensure that measurements made by each instrument overlap, a restriction was placed on ACE-FTS and MIPAS measurements: that their retrieved vertical profiles extend to low enough altitudes, after applying data quality criteria. For ACE-FTS, this requirement was 10 km. For MIPAS, this requirement was relaxed to less than 12 km. IMK-IAA MIPAS CH_4 VMR vertical profile retrievals do not extend as low as those made by ESA, to the extent that having the same restriction on altitude range results in only a quarter as many coincidences as the ESA data product. Relaxing the constraint to only 12 km maintains the assurance that retrieved VMRs will overlap with the TANSO-FTS altitude range, though there are only 60 % as many IMK-IAA coincidences compared to ESA coincidences.

TANSO-FTS makes nadir observations in a grid pattern by sweeping its line-of-sight across its ground-track. This results in a high density of vertical profiles, such that, for a single observation made by ACE-FTS, MIPAS, or NDACC, there are an average of 11 coincident TANSO-FTS measurements. The subsequent measurement made by MIPAS or an NDACC station will be coincident with a similar number of TANSO-FTS measurements, and most of those will also be coincident with the previous MIPAS or NDACC measurement. A common way to deal with multiple coincidences is to take the mean of the VMR vertical profiles from each instrument, and to compute the difference of the means (e.g., Holl et al., 2016) . When comparing MIPAS to TANSO-FTS, however, this results in some measurements contributing to the analysis more times than others, biasing the computed VMR difference profiles. Furthermore, this leads to using a mean TANSO-FTS VMR vertical profile that is strongly smoothed, while a coincident ACE-FTS (or NDACC, depending on the station's rate of acquisition at the time) VMR vertical profile is not.

To reduce biases caused by over-counting, when comparing TANSO-FTS to MIPAS, and by smoothing, when comparing TANSO-FTS to ACE-FTS, we reduced the number of coincident measurements by seeking a set of one-to-one coincidences for unique measurements in the sparser ~~dataset~~ data set (which is always ACE-FTS, MIPAS, or NDACC). For each measurement that is being compared to TANSO-FTS, we find the TANSO-FTS measurement with the minimum of the sum of ratios of distance in space and time to the coincidence criteria, giving equal weight to both parameters as: $\min(dx/x_{crit} + dt/t_{crit})$, where dx and dt are the distance and time between a given measurement and a TANSO-FTS coincidence, and x_{crit} and t_{crit} are the coincidence criteria. This method is similar to using a ~~z-~~standard score to compare the spatial and temporal separation, but the sample size of the set of TANSO-FTS measurements coincident with another measurement is on the order of only ten. Furthermore, the mean and standard deviations of dx and dt reflect the time and distance between each consecutive TANSO-FTS measurement, rather than the time and spatial separation between each TANSO-FTS measurement and those from MIPAS, ACE-FTS, or NDACC.

Table 2 shows the total number of coincidences found between TANSO-FTS and each validation target instrument, as well as the subsets of unique TANSO-FTS measurements and the one-to-one coincidences used in this paper (equivalent to the number of unique measurements made by each target instrument). Fig. 2 shows an example of the global distribution of coincident measurements. Shown are the first 200 one-to-one coincidences after 1 January 2012. For the ESA and IMK-IAA MIPAS data products, this number of coincidences is found in around two weeks. For ACE-FTS and the NDACC stations (combined), these coincidences occur over several months.

5 Averaging kernels

The averaging kernels of a profile retrieval provide information about the contributions of the retrieval from a priori information and the measurements. In this study, the retrieval methods for each data set differ, and the averaging kernel matrices are differently defined. In general, the rows of the averaging kernel matrix are peaked functions ~~with widths indicative of~~ whose full-width at half maximum (FWHM) can be used to define the vertical resolution of the measurement. The sum of the rows of the matrix gives the sensitivity, or response, of the retrieval. A sensitivity close to one indicates that most of the information in the retrieval comes from the measurement, while sensitivities less than one indicate increased reliance on the a priori in the solution.

The rows of the averaging kernel matrices for the ESA MIPAS, IMK-IAA MIPAS, TANSO-FTS, and the Eureka NDACC station are shown in Fig. 3. Each panel shows the mean from 30 retrievals~~, with each averaging kernel interpolated to~~. Vertical profiles of pressure associated with each retrieval's averaging kernel matrix are, in general, unique, so a common pressure grid ~~for that instrument~~ was selected for each instrument and averaging kernels were interpolated prior to averaging.

In this study, we treat TANSO-FTS retrievals as having the coarser vertical resolution in all cases, despite the widths of the kernel functions shown in Fig. 3a, which are comparable to MIPAS and narrower than NDACC. The peak locations of the TANSO-FTS averaging kernels do not match the corresponding pressure level of each kernel~~, therefore~~. Therefore the full-width at the half-maximum values when considering the location of the appropriate pressure level are much larger than the full-width at half-maximum values for the averaging kernels of the other instruments.

In the NDACC retrievals, the a priori has a large role, and information coming from the measurements can hardly distinguish the contribution coming from the different altitudes. This leads to wide, overlapping averaging kernels. The IMK-IAA MIPAS retrievals use a form of Tikhonov regularization without an a priori. The ESA MIPAS retrievals use the regularizing Levenberg-Marquardt approach (where the parameter setting has been chosen to leave results largely independent from the initial guess profiles) and a posteriori Tikhonov regularization without an a priori. The ACE-FTS retrievals do not use a regularized matrix inverse method. Consequently, the ACE-FTS and IMK-IAA MIPAS averaging kernels are very narrow, their peak values are close to one at each altitude where a spectrum was acquired, and the solutions do not rely on a priori information. Very similar averaging kernel are obtained also for ESA MIPAS, with wider widths at lower altitudes where the retrieval grid used is coarser than the measurement grid. The sensitivity of both ACE-FTS and MIPAS, shown in Fig. 3e, is close to one at all altitudes, falling off above 60 or 70 km. ~~This is shown in Fig. 3e.~~ ACE-FTS averaging kernels are under development, and preliminary work is shown in Sheese et al. (2016).

The typical sensitivity of an NDACC retrieval is close to unity until above 20 km, falling off towards zero through 60 km. The sensitivity of TANSO-FTS only reaches 0.2–0.3 between 5–10 km. The implication of such low values for sensitivity is that the TANSO-FTS retrievals are highly dependant on their a priori.

The trace of the averaging kernel matrix gives the DOFs~~, which are shown for the~~. For example, DOFs for retrievals made by TANSO-FTS, IMK-IAA MIPAS, ESA MIPAS, and NDACC ~~retrievals are shown in Fig. 4 for retrievals made in the Arctic(for example)~~from observations over the Arctic, above 60° N~~.~~, are shown in Fig. 4. The IMK-IAA MIPAS and TANSO-FTS data are in coincidence with one another. The NDACC data come from Eureka, Ny Ålesund, and Thule. The NDACC and ESA MIPAS data shown are the TANSO-FTS one-to-one coincidences used throughout this study (but are not coincident with the TANSO-FTS data shown in the top panel of Fig. 4). The trends visible are seasonal and are related to opacity and water vapour content. Recreating this figure over mid-latitudes or the tropics reveals a flat trend over time, while over Antarctica, the trends are reversed in DOFs-space.

The mean of the DOFs for the three NDACC stations over the Arctic is 1.98 with a standard deviation, σ, of 0.50. Over the tropics, considering data from Izaña, La Réunion St. Denis, Altzomoni, and Mauna Loa (La Réunion Maïdo only has data from 2013 onward, not shown here), the mean is 2.39 with $\sigma = 0.37$. The mean DOFs for IMK-IAA MIPAS are slightly larger than those for ESA MIPAS. Over the Arctic, their means and standard deviations are 17.05, $\sigma = 1.06$ and 15.76, $\sigma = 0.93$, for IMK-IAA and ESA, respectively. Over the tropics, they are 16.10, $\sigma = 0.33$ and 15.88, $\sigma = 1.20$.

The TANSO-FTS DOFs are larger at low latitudes, with a mean over the tropics of 0.72 and $\sigma = 0.08$, and means over the Arctic and Antarctic of 0.32 and 0.20, respectively ($\sigma = 0.13$ and 0.12). The DOFs for a TANSO-FTS retrieval rarely go above unity. Conversely, in the coincident NDACC data discussed above, over the tropics and Arctic, the DOFs never fall below unity. Note that the averaging kernel matrices for TANSO-FTS, and therefore the DOFs, cover a much smaller altitude range than for NDACC and MIPAS, which can extend above 100 km.

6 VMR vertical profile comparisons

6.1 Methodology

Retrievals made by an instrument with fine vertical resolution may result in structure over its vertical range that is not distinguishable in retrievals made by an instrument with coarser vertical resolution. In order to make the best comparison between two instruments with differing vertical resolution, it is necessary to smooth the vertical profiles retrieved from the finer resolution instrument, in order to simulate what we could infer from it if it had a similar sensitivity as the other instrument. Smoothing is done using the a priori $\mathrm{CH_4}$ VMR vertical profiles and averaging kernel matrices of the instrument with lower vertical resolution (Rodgers and Connor, 2003):

$$\hat{\mathbf{x}}_s = \mathbf{x}_a + \mathbf{A}(\hat{\mathbf{x}} - \mathbf{x}_a), \tag{1}$$

where $\hat{\mathbf{x}}$ is original higher-resolution retrieved profile, $\hat{\mathbf{x}}_s$ is the smoothed profile, $\mathbf{x}_a$ is the a priori profile of the lower-resolution retrieval, and $\mathbf{A}$ is the averaging kernel matrix of the lower-resolution retrieval. $\mathbf{x}_a$ and $\mathbf{A}$ are from the TANSO-FTS retrieval in all cases presented here. The smoothed profile, $\hat{\mathbf{x}}_s$, approximates the a priori, $\mathbf{x}_a$, when either the rows of $\mathbf{A}$ are close to zero, or when the retrieval is close to $\mathbf{x}_a$. As can be inferred from Fig. 3a, above 20–25 km $\hat{\mathbf{x}}_s \sim \mathbf{x}_a$.

In order to apply Eq. 1, all the variables on the right hand side must be interpolated to a common grid. TANSO-FTS retrievals are done on a retrieved pressure grid. Determining the altitude of its VMR vertical profiles requires applying the equation of hydrostatic equilibrium, and incorporating a priori temperature and water vapour. Since pressure is retrieved by ACE-FTS and MIPAS, and the tropospheric a priori pressure profiles and measured surface pressure are accurate for NDACC Sepúlveda et al. (2014) , all comparisons here have been done on a common pressure grid, as opposed to an altitude grid.

The data products do not always overlap over the entire pressure range of the common grid. Extrapolation is needed to ensure that the length of $\hat{\mathbf{x}}$ matches the dimensions of $\mathbf{A}$ in Eq. 1. For ACE-FTS and MIPAS, we use $\mathbf{x}_a$ to extend their retrieved profiles below their altitude range to cover the full pressure range of the TANSO-FTS averaging kernels. The averaging kernels at these non-overlapping pressure levels do not contribute to the smoothed retrieval at higher, overlapping levels. The following steps are taken to compute vertical profiles of the mean CH_4 VMR differences:

1. appropriate instrument data quality flags are applied to each VMR vertical profile in the coincidence pair,

2. TANSO-FTS a priori and validation target VMR vertical profiles are interpolated to the TANSO-FTS retrieval pressure grid,

3. the interpolated validation target profile is extended as needed to match the TANSO-FTS pressure range (and vector length) using the TANSO-FTS a priori,

4. the interpolated validation target profile is smoothed using the TANSO-FTS averaging kernel matrix using Eq. 1,

5. TANSO-FTS retrieved and validation target smoothed VMR vertical profiles are interpolated to a standard pressure grid, levels outside the pressure range of the target's VMR profile are discarded,

6. the piecewise difference between the TANSO-FTS and the smoothed validation target VMR vertical profiles is found,

7. the means, standard deviations, and correlation coefficients of the VMR differences are calculated at each level of the standard pressure grid for all coincidences within a latitude zone.

For comparison, mean VMR vertical profile differences were also computed without smoothing by using only steps 1, 5, 6, and 7. Zonally averaged VMR difference profiles are presented in Sect. 6.2 and results obtained without applying smoothing to the validation targets are shown in Sect. 6.3. The data quality flags in step 1, referring to variables in the data product files, were, for TANSO-FTS: $CH4ProfileQualityFlag$ must be zero; for ACE-FTS: $quality\_flag$ must be zero, and cannot be equal to four, five, or six at any altitude; for ESA MIPAS: $ch4\_vmr\_validity$ must be one and $pressure\_error$ cannot be NaN; for IMK MIPAS: $visibility$ must be one, and $akm\_diagonal$ must be greater than 0.03.

Holl et al. (2016) found that ~~looking for~~ identifying and removing coincident CH_4 VMR vertical profile pairs that may have one or both profile locations within a polar vortex, and then filtering these events, had little effect on their vertical profile comparisons below 25 km. Polar vortex event will have a much smaller ~~affect~~ effect on this study since it uses global and year-round data sets. For these two reasons, our method does not filter for profiles located within a polar vortex. Arrival Heights may be differently affected by a much-stronger Antarctic polar vortex, but comparison results from this site are not anomalous and only accounts for 1.5 % of the NDACC data set so are treated in a consistent manner.

6.2 Zonally averaged VMR profile differences

Following Holl et al. (2016), we are trying to determine whether there are any zonal biases in the TANSO-FTS data, or zonal dependencies when making comparisons to other instruments. The mean CH_4 VMR differences, averaged zonally, between the TANSO-FTS vertical profiles and the smoothed vertical profiles from ACE-FTS, IMK-IAA MIPAS, ESA MIPAS, and each NDACC station are show in Fig. 5. Each row in Fig. 5 shows the results from five latitudinal zones: 90–60° N, 60–30° N, 30° N–30° S, 30–60° S, and 60–90° S. The left-most column shows the mean differences between the retrievals from TANSO-FTS and those from the other instruments, always calculated as TANSO-FTS $- target$. One standard deviation is shown for each instrument comparison with dotted lines. The middle-left column shows the mean differences as a percentage of the mean CH_4 VMR vertical profile taken for the target validation instrument in each zone. The number of VMR measurements used in the mean at each altitude, for each comparison, is shown in the right-most panel, with ESA MIPAS always having the most. At each altitude, we also calculated the Pearson correlation coefficient between the set of TANSO-FTS CH_4 VMR measurements and the coincident set from each validation instrument. These are shown in the middle-right column for each panel in Fig. 5.

For each zone, the mean difference tends towards zero and the standard deviation falls off above 100 hPa. This is a reflection of the TANSO-FTS sensitivity. Above this altitude, the TANSO-FTS averaging kernels tend to zero, as shown in Fig. 3, and the smoothed profiles from each target instrument begin to approximate the TANSO-FTS a priori. Likewise, the TANSO-FTS retrieval above this pressure level is also close to its a priori. Conversely, the number of CH_4 VMR measurements in the mean falls off sharply below 10–12 km, or around 80–90 hPa, for the comparisons to the satellite instruments. For the satellite instruments and many of the NDACC stations we see the same trend: a positive bias (TANSO-FTS VMRs are greater than those of the validation instruments) decreasing with increasing altitude, with a tropospheric mean of around ~~0.02~~ 20 ppbv, or 1 %. The bias is smallest for the two MIPAS data products in the tropics, between 30° N and 30° S. The bias relative to ACE-FTS is consistent in all the zones. For three of the NDACC stations, Ny Ålesund, Bremen, and Toronto, there is a negative bias (TANSO-FTS retrieves less CH_4 than these stations), and for Eureka and Jungfraujoch the bias is close to zero.

There is a notable feature just below 100 hPa in all the zones except 30–60° S. This feature is a pronounced increase in the mean difference in the northern zones 60–30° N and 90–60° N, while it is a decrease in the mean difference between 30° N–30° S, and 60–90° S. It is around this pressure level, or altitude, that the VMR of CH_4 begins to fall off rapidly from between 1.8 to 2 ppmv in the troposphere towards 0 ppmv in the upper stratosphere and mesosphere. This feature indicates that the altitude at which this VMR decrease occurs differs between instruments. In the northern hemisphere, this decrease in CH_4 VMR occurs at higher altitudes for TANSO-FTS than for the other instruments, and in the tropics and southern hemisphere, this decrease occurs more rapidly and at lower altitudes for TANSO-FTS.

For all instruments and in all zones, the correlation coefficients, R^2, at each altitude fall off very sharply, to around 0.2, below the 90 hPa level (and remain higher in the tropics). This indicates that biases seen in the mean differences are not uniform across the coincident data set and that there is significant variability in the magnitudes of the differences for individual vertical profile pairs, and in the direction of the difference. This is related to the increasing standard deviation of the differences with decreasing altitude, but also to the standard deviations of each data product in the comparison. The sharpness and altitude of the decrease is directly related to the TANSO-FTS averaging kernels. Above the 100 hPa level, the standard deviations of the TANSO-FTS and the smoothed validation target fall off very sharply as they both begin to approximate the a priori (which also explains why R^2 is close to 1).

6.3 Impact of smoothing

This study was also performed without applying any smoothing to the vertical profiles of the target validation instruments. These results are shown in Fig. 6, which has the same panels as Fig. 5. The data have not been separated zonally, and the plots show means for all latitudes. No zonal biases were observed in the unsmoothed data. The 16 NDACC stations have been combined into a single data set.

~~This study reveals the actual differences one would expect when using~~ Fig. 6 shows the mean differences between the TANSO-FTS data product and those of other instruments, and the behaviour of the comparisons at higher altitudes when the validation targets are unaffected by the TANSO-FTS averaging kernels. Without the smoothing applied, the ~~differences~~ difference profiles in Fig. 6 show more consistent behaviour over the pressure, or altitude, range shown. While the magnitude of the differences is much greater without smoothing, it is not consistently biased high or low for all the data products at all altitudes. When comparing to the satellite instruments in the upper troposphere, we find that the TANSO-FTS retrieval has greater CH_4 VMRs by around ~~0.05~~50 ppbv, or around 3 %.

For context, a comparison between the ACE-FTS and ESA MIPAS data products, using profiles that were coincident with the same TANSO-FTS observation, is shown in grey. The mean differences between these two data products are smaller than those relative to TANSO-FTS, but have comparable standard deviations, and a slightly smaller correlation, with $R^2 = 0.5$ and 0.6 in the upper troposphere.

The comparison between TANSO-FTS and NDACC extends below the range of ACE-FTS and MIPAS. NDACC and TANSO-FTS agree very well in this region, between ~~± 0.03~~± 30 ppbv, or between ± 2 %. In this case, the NDACC stations retrieve more CH_4, on average. The low-altitude NDACC and TANSO-FTS data are also more closely linearly correlated, between 50 and 60 %. It should also be noted that the standard deviation of the TANSO-FTS and NDACC differences is also less than those for ACE-FTS and MIPAS at all altitudes.

7 Partial column comparisons

7.1 Methodology

For each CH_4 VMR vertical profile in a pair of coincident measurements, we computed a partial column and compared those from TANSO-FTS to each of the other instruments to investigate how well correlated the derived CH_4 abundances are. For consistency, each pair of partial columns must be calculated over the same pressure range, as the number of molecules in the column strongly depends on the altitude range (length of the column) of the integral. To determine the pressure range over which to compute partial columns for each coincident pair of profiles, we considered the TANSO-FTS averaging kernels.

We investigated the sensitivity of the TANSO-FTS retrievals, as defined in Sect. 5 to find an altitude range which minimizes the partial column dependence on a priori information, ensuring our investigation is focused on retrieved information from TANSO-FTS. Fig. 7 shows a two dimensional histogram of the number of TANSO-FTS profiles, for all validation targets combined for two criteria: setting a requirement that the sensitivity must be greater than some threshold, and the resulting number of usable pressure levels in the integral for each profile. We see that the maximum number of usable levels falls off in an approximately linear manner with increasing sensitivity threshold, and that for any sensitivity threshold there will be a large number of TANSO-FTS CH_4 VMR vertical profiles that never meet the sensitivity criteria. Increasing the sensitivity cutoff by 0.05 causes approximately 10,000 additional TANSO-FTS vertical profiles, or around 6 % of the total data set combining all validation targets, to fail to meet the requirement at any altitude. The number of usable pressure levels given a restriction on sensitivity is not normally distributed, as can be inferred from the empty area in the upper right of Fig. 7.

For this study, we have selected a sensitivity threshold of 0.2 and require a minimum of three integrable pressure levels. Approximately 23 % of the TANSO-FTS retrievals do not meet these criteria. In such a case, partial columns are still computed using three pressure levels surrounding the level with the maximum sensitivity that are within the range of the target profile (e.g., not below 10 km when comparing to ACE-FTS). These excluded data do not exhibit a broader distribution, but their computed partial columns are all very small due to the integration range. Because the overlapping altitude regions for NDACC and TANSO-FTS measurements extend much lower in the atmosphere than for ACE-FTS and MIPAS, the number of TANSO-FTS profiles that do not meet the sensitivity criteria is much smaller for NDACC.

Partial columns are computed as:

$$\text{Column} = \int_{z_1}^{z_2} \frac{P(z)}{kT(z)} \chi(z) dz, \quad (2)$$

where z_1 and z_2 bound the integration range over altitude z, P is pressure, T is temperature, χ is the CH_4 VMR, and k is the Boltzmann constant. For each instrument, $\chi(z)$ is the retrieved quantity, and retrievals were either performed on a pressure grid, or pressure was retrieved simultaneously. We compute partial columns from vertical profiles after step 5 in Sect. 6.1, so both the TANSO-FTS and the smoothed validation target profiles have the same pressure at each level in the integration. Since TANSO-FTS retrievals do not have an altitude grid, we use that of the coincident measurement, which corresponds to the pressure levels and should be very accurate within the altitude range considered in this study (upper troposphere to lower stratosphere). Thus, we are integrating over the same altitude range for both instruments. Since ACE-FTS and both MIPAS data products include retrieved temperatures, we use their retrieved temperature. For TANSO-FTS and NDACC, we use their corresponding a priori temperatures.

Several methods of integration were investigated and the results presented in Sect. 7.2 are derived by simple summation of the integrand multiplied by the bin-width of each data point in km. We also used numerical integration techniques, variations of Newton-Coates and Gaussian quadrature formulas. These did not provide significantly different results due the large size of our sample (i.e., our results are statistics found from the Least-squares method, and small differences in the individual partial columns due to different integration methods do not introduce bias). Since the analytic function being integrated is not well defined, neither is the uncertainty of the derived partial column. Propagating reported retrieval uncertainties of temperature and VMR provides the most appropriate estimate of uncertainty, which is shown in Fig. 8.

7.2 Partial column correlation

The computed partial columns from TANSO-FTS are plotted against of those from each validation instrument in Fig. 8. The panels for ACE-FTS, ESA MIPAS, and IMK-IAA MIPAS contain measurements for all latitudes, and that for NDACC combines results from all 16 stations. Since IMK-IAA retrievals do not extend as low as those of ESA generally, the altitude range of the partial column integral is often smaller than those of the other instruments, resulting in smaller CH_4 abundances. Conversely, abundances when comparing to the NDACC stations are the largest.

The Pearson correlation coefficients, R^2, are: 0.9986, ~~0.9968,~~ 0.9965, ~~and 0.9929~~ 0.9968, and 0.9958 for ACE-FTS, ~~ESA MIPAS,~~ IMK-IAA MIPAS, ESA MIPAS and NDACC, respectively. The slopes of the fitted correlation lines are all close to unity, and a small bias is seen in the y-intercept corresponding to between 0.4 % and 2.8 % relative to the mean partial columns of the validation targets, with the greatest corresponding to the NDACC data. Among the individual NDACC stations, those with the largest correlation function intercept are Mauna Loa, Jungfraujoch, Bremen, Izaña, and Zugspitze (1.2×10^{23}–7.5×10^{23}). TANSO-FTS has a negative intercept only with respect to two stations: The correlation coefficients for each station are all greater than 0.96, except for Mauna Loa, Izaña, and Maïdo, La Réunion, which all happen to be islands, and for which a large number of coincident TANSO-FTS measurements would have been made over water (see Sect. 8).

Statistics regarding the distribution of the integration ranges over altitude are given in Table 3. This table gives the number of coincident pairs for each validation instrument for which the TANSO-FTS CH_4 VMR vertical profile passed the sensitivity requirements. It also gives the mean and standard deviation of the lower bound of the integral (lower altitude), the width of the interval (highest altitude minus the lowest), and the number of pressure levels used. As expected, the NDACC stations have the widest altitude range, while the IMK-IAA MIPAS retrievals have the smallest. Note that the column in Table 3 showing number of levels used does not correspond to the mode in Fig. 7 since Fig. 7 considers only the TANSO-FTS averaging kernels and does not reflect the lack of available comparison data at lower altitudes.

Repeating the analysis using unsmoothed data from ACE-FTS, ESA and IMK-IAA MIPAS, and NDACC, the spread in the correlation plots increases and the biases observed in the intercepts increase, while the correlation coefficients remain very close to unity. Fig. 9 shows derived partial column correlation plots for each validation target instrument. The intercept, without smoothing is between 2 and 6 %. The correlation coefficient for the MIPAS instruments is reduced to 0.97.

8 Discussion

The objective of this study was to quantitatively assess TANSO-FTS CH_4 VMR vertical profile retrievals compared with other FTS instruments, and to further investigate whether there were any biases with latitude or other retrieval parameters. As shown in Sect. 6.2, we did not find a significant difference in mean CH_4 VMR profile differences between latitudinal zones.

To investigate further, we consider the CH_4 VMR differences averaged over altitude for each coincident pair, for each validation instrument. To choose the altitude range over which to find the mean, we use the same sensitivity criteria developed in Sect. 7.2. The resulting mean differences between TANSO-FTS and ACE-FTS, MIPAS, and NDACC are shown as a function of latitude in Fig. 10. ~~A bias is seen~~ Weighted least squares regression of the combined data sets for each hemisphere reveals a bias at all latitudes of ~~0.01330 ± 0.00006~~13.30 ± 0.06 ~~, when combining results from all four validation instruments~~ppbv. There is also a small slope in the data from each hemisphere, decreasing from the poles to the tropics. Linear fit parameters for the combined data ~~set~~ sets in each hemisphere are given in Table 4. This leads to a bias of around ~~0.004~~4 ppbv in the tropics (0.25 % of a tropical tropospheric VMR value of 1.8–2 ppmv), and of 0.014 ppmv ~~or~~ and 0.020 ppmv at the North and South Pole, respectively (or around 1 %). The biases are latitude-dependent and vary between the tropics and the poles.

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). 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. We found no biases in our coincident TANSO-FTS ~~dataset~~ data set related to any of these parameters, or whether the observation was made during night or day. The land or sea mask is an indicator of whether the retrieval was made over land, water, or a combination in the field-of-view. In our data set of all one-to-one coincidences between TANSO-FTS and the validation targets, 54.0 % of TANSO-FTS measurements were made over water, 36.3 % were made over land, and 9.6 % were a mixture. The sunglint flag indicates whether the positions of the sun, satellite, and observation point are related within a predefined range, qualifying the observation as being made in sun-glint mode. In our data set, only 1.6 % of TANSO-FTS measurements are sun-glint observations, and they are all over water and between ±45° latitude. Finally, 54.1 % of TIR observations were made at night.

~~We have investigated the sensitivity, and averaging kernels for the~~ 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 CO_2, increasing their sensitivity and ability to accurately constrain CH_4 ~~TIR~~ 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 CH_4 vertical profiles from different air masses. Our investigation of the NDACC data provides an estimate of the dependence of the 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, e.g., Boone et al. (2013); Rothman et al. (2013); Toon et al. (2016) . 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 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.

9 Conclusions

The TANSO-FTS TIR CH_4 vertical profile data product , and 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. We have investigated the sensitivity and averaging kernels for the TANSO-FTS data product, and done a global comparison with four other FTS data products. Our comparisons showed that the sensitivity of the TANSO-FTS retrieval is relatively low at all altitudes, and that there is a limitation on the useful upper altitude of its data product of below around 15 or 20 km. However, this vertical extent is below Unfortunately, the lower altitude boundaries (10–15) of the other satellite-based data products, between 7–15 km, reduces the vertical range over which we can make comparisons. In the upper troposphere, we found good agreement between TANSO-FTS and NDACC, without a bias. The agreement between these two data sets persisted regardless of whether smoothing was applied to the NDACC data. Therefore, despite the lower sensitivity of the TANSO-FTS data product, it remains an important and unique data set of global tropospheric CH_4 measurements.

In the overlapping altitude ranges of the three satellite data products, we found a small, but consistent, positive bias of around 0.0220 ppbv, or 1 %. We found that the shape of the profile shapes of the TANSO-FTS CH_4 VMR vertical profiles near 15 km, where the CH_4 VMR vertical profiles fall falls off with increasing altitude, does not match that 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. Despite the large variability in each data set and in the differences between the TANSO-FTS retrievals and the others, we found that partial columns computed from the vertical profiles were very tightly correlated, with and without smoothing.

When looking for a relationship between latitude and the differences between data products, we found a small, but statistically significant, dependence of the mean differences , taken over altitude and vertically-averaged differences on latitude. The TANSO-FTS data product shows better agreement over the tropics than the poles.

We look forward to future versions of the retrieval which may feature a greater sensitivity and altitude range, while reducing the small biases and dependence on the a priori profiles. 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.

Author contributions. The work presented here was done by K. S. Olsen with input from, and under the supervision of K. Strong, K. A. Walker, and N. Saitoh. K. S. Olsen prepared the manuscript with contributions from all co-authors. CH_4 retrievals were developed and provided, with additional input, by: N. Saitoh for TANSO-FTS, C. D. Boone for ACE-FTS, P. Raspollini for ESA MIPAS, J. Plieninger for IMK-IAA MIPAS, M. Grutter for Altzomoni, J. W. Hannigan for Thule and Mauna Loa, F. Hase for Kiruna, N. Jones for Wollongong, W. Bader for Jungfraujoch, M. de Mazière for La Réunion St. Denis and Maïdo, J. Notholt for Bremen and Ny Ålesund, M. Schneider for Izaña, D. Smale for Lauder and Arrival Heights, S. Conway and K. Strong for Eureka and Toronto, and R. Sussmann for Zugspitze.

Acknowledgements. This research was conducted under the framework of the Japan Aerospace Exploration Agency (JAXA), National Institute for Environmental Studies (NIES), and the Ministry of the Environment (MOE) Research Announcement (RA) project “GOSAT Validation Using Eureka and Toronto Ground-Based Measurements and ACE, CloudSat, and CALIPSO Satellite Data” with Kimberly Strong (Principal Investigator,University of Toronto), David Hudak (Co-Investigator, Environment Climate Change Canada (ECCC)), N. T. O’Neill (Co-Investigator, Universitde Sherbrooke), and K. A. Walker (Co-Investigator, University of Toronto). This GOSAT RA project ,” which was supported by the Canadian Space Agency (CSA), the Natural Sciences and Engineering Research Council of Canada (NSERC), and ECCCEnvironment & Climate Change Canada (ECCC). This work is the result of many , long-lasting collaborations, and we would like to thank our co-authors and collaborators for providing data and expertise. The GOSAT team provided early access to its TANSO-FTS TIR CH_4 VMR vertical profiles and regular access to its other data products through the GOSAT User Interface Gateway: data.gosat.nies.go.jp/. SCISAT/ACE is a Canadian-led mission mainly supported by the CSA and NSERC. The ACE-FTS Science Team at the University of Waterloo provided access to their level 2 data through databace.scisat.ca/level2/ace_v3.5, and expert knowledge with its interpretation and quality management. Access to the ESA MIPAS level 2 data was granted and provided through the ESA Earth Online portal: earth.esa.int. The IMK-IAA MIPAS level 2 data were accessed using KIT website: www.imk-asf.kit.edu/english/308.php. NDACC data has been compiled by many independent research groups, and was accessed through a National Centers for Environmental Prediction FTP server, with each station being accessible through: www.ndsc.ncep.noaa.gov/data/data_tbl/. Measurements at PEARL were made by the Canadian Network for the Detection of Atmospheric Change (CANDAC), led by J. R. Drummond, and in part by the Canadian Arctic ACE/OSIRIS Validation Campaigns, led by K. A. Walker. Support is provided by AIF/NSRIT, CFI, CFCAS, CSA, EC, GOC-IPY, NSERC, NSTP, OIT, PCSP and ORF. Logistical and operational support is provided by PEARL Site Manager P. Fogal, CANDAC operators, and the ECCC Weather Station. Measurements at the University of Toronto Atmospheric Observatory were supported by CFCAS, ABB Bomem, CFI, CSA, EC, NSERC, ORDCF, PREA, and the University of Toronto. NDACC data analysis at Toronto and Eureka was supported by the CAFTON project, funded by the CSA’s FAST Program. The National Institute of Water and Atmospheric Research Ltd (NIWA) operated ground based ground-based

FTSs are core-funded through New Zealand's Ministry of Business, Innovation and Employment. We thank Antarctica, New Zealand and the Scott Base staff for providing logistical support at Arrival Heights. Measurements at the Jungfraujoch station are primarily supported by the Fonds de la Recherche Scientifique (F.R.S.–FNRS) and the Fédération Wallonie-Bruxelles, both in Brussels. The Swiss GAW-CH program of MeteoSwiss is further acknowledged. We thank the International Foundation High Altitude Research Stations Jungfraujoch and Gornergrat (HFSJG, Bern) for supporting the facilities needed to perform the observations and the many colleagues who contributed to FTS data acquisition at that site. W. Bader has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 704951. We would like to thank A. Bezanilla, who operates the Altzomoni site, and W. Stremme who does the data processing for the Altzomoni site and uploads the data to the ~~NDACCW~~ NDACC archive. Altzomoni is supported by Consejo Nacional de Ciencia y Tecnología (CONACYT, grants 239618 & 249374) and la Dirección General de Asuntos del Personal Académico de la Universidad Nacional Autónoma de México (DGAPA-UNAM, grants IN109914 & IA101814).

References

Bader, W., Bovy, B., Conway, S., Strong, K., Smale, D., Turner, A. J., Blumenstock, T., Boone, C., Collaud Coen, M., Coulon, A., Garcia, O., Griffith, D. W. T., Hase, F., Hausmann, P., Jones, N., Krummel, P., Murata, I., Morino, I., Nakajima, H., O'Doherty, S., Paton-Walsh, C., Robinson, J., Sandrin, R., Schneider, M., Servais, C., Sussmann, R., and Mahieu, E.: The recent increase of atmospheric methane from 10 years of ground-based NDACC FTIR observations since 2005, Atmos. Chem. Phys., 17, 2255–2277, https://doi.org/10.5194/acp-17-2255-2017, 2017.

Baray, J. L., Courcoux, Y., Keckhut, P., Portafaix, T., Tulet, P., Cammas, J. P., Hauchecorne, A., Godin, S., Beekmann, S., de Mazière, Hermans, C., Desmet, F., Sellegri, K., Colomb, A., Ramonet, M., Sciare, J., Vuillemin, C., Hoareau, C., Dionisi, D., Duflot, V., Vérèmes, H., Porteneuve, J., Gabarrot, F., Gaudo, T., Metzger, J.-M., Payen, G., Leclair de Bellevue, J., Barthe, C., Posny, F., Ricaud, P., Abchiche, A., and Delmas, R.: Maïdo observatory: a new high-altitude station facility at Reunion Island (21° S, 55° E) for long-term atmospheric remote sensing and in situ measurements, Atmos. Meas. Tech., 6, 2865–2877, https://doi.org/10.5194/amt-6-2865-2013, 2013.

Batchelor, R. L., Strong, K., Lindenmaier, R., Mittermeier, R. L., Fast, H., Drummond, J. R., and Fogal, P. F.: A new Bruker IFS 125HR FTIR spectrometer for the Polar Environment Atmospheric Research Laboratory at Eureka, Canada: measurements and comparison with the existing Bomem DA8 spectrometer, J. Atmos. Ocean. Tech., 26, 1328–1340, https://doi.org/10.1175/2009JTECHA1215.1, 2009.

Baylon, J. L., Stremme, W., Plaza, E., Bezanilla, A., Grutter, M., Hase, F., and Blumenstock, T.: CO_2 total column variability from ground-based FTIR measurements over central Mexico, in: AGU Fall Meeting, AGU Fall Meeting, 2014.

Bernath, P. F.: The Atmospheric Chemistry Experiment (ACE), J. Quant. Spectrosc. Radiat. Transfer, 186, 3–16, https://doi.org/10.1016/j.jqsrt.2016.04.006, 2017.

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P.-F., Colin, R., DeCola, P., de Mazière, M., Drummond, J. R., Dufour, D., Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn, E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F., Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M.-A., Strong, K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J.: Atmospheric Chemistry Experiment (ACE): Mission overview, Geophys. Res. Lett., 32, L15S01, https://doi.org/10.1029/2005GL022386, 2005.

Blumenstock, T., Kopp, G., Hase, F., Hochschild, G., Mikuteit, S., Raffalski, U., and Ruhnke, R.: Observation of unusual chlorine activation by ground-based infrared and microwave spectroscopy in the late Arctic winter 2000/01, Atmos. Chem. Phys., 6, 897–905, https://doi.org/10.5194/acp-6-897-2006, 2006.

Boone, C. D., Nassar, R., Walker, K. A., Rochon, Y., McLeod, S. D., Rinsland, C. P., and Bernath, P. F.: Retrievals for the atmospheric chemistry experiment Fourier-transform spectrometer, Appl. Opt., 44, 7218–7231, https://doi.org/10.1364/AO.44.007218, 2005.

Boone, C. D., Walker, K. A., and Bernath, P. F.: Version 3 Retrievals for the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS), in: ACE at 10: Solar Occultation Anthology, edited by Bernath, P. F., pp. 103–127, A. Deepak Publishing, Hampton, Virg., 2013.

Buchwitz, M., Schneising, O., Burrows, J. P., Bovensmann, H., Reuter, M., and Notholt, J.: First direct observation of the atmospheric CO_2 year-to-year increase from space, Atmos. Chem. Phys., 7, 4249–4256, https://doi.org/10.5194/acp-7-4249-2007, 2007.

Côté, J. S., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The operational CMC-MRB global environmental multiscale (GEM) model. Part I: Design considerations and formulation, Mon. Weather Rev., 126, 1373–1395, https://doi.org/10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2, 1998.

de Mazière, M., Vigouroux, C., Bernath, P. F., Baron, P., Blumenstock, T., Boone, C., Brogniez, C., Catoire, V., Coffey, M., Duchatelet, P., Griffith, D., Hannigan, J., Kasai, Y., Kramer, I., Jones, N., Mahieu, E., Manney, G. L., Piccolo, C., Randall, C., Robert, C., Senten, C., Strong, K., Taylor, J., Tétard, C., Walker, K. A., and Wood, S.: Validation of ACE-FTS v2.2 methane profiles from the upper troposphere to the lower mesosphere, Atmos. Chem. Phys., 8, 2421–2435, https://doi.org/doi:10.5194/acp-8-2421-2008, 2008.

Engel, A., Bönisch, H., Schwarzenberger, T., Haase, H.-P., Grunow, K., Abalichin, J., and Sala, S.: Long-term validation of ESA operational retrieval (version 6.0) of MIPAS Envisat vertical profiles of methane, nitrous oxide, CFC11, and CFC12 using balloon-borne observations and trajectory matching, Atmos. Meas. Tech., 9, 1051–1062, https://doi.org/10.5194/amt-9-1051-2016, 2016.

Errera, Q., Ceccherini, S., Christophe, Y., Chabrillat, S., Hegglin, M. I., Lambert, A., Ménard, R., Raspollini, P., Skachko, S., van Weele, M., and Walker, K. A.: Harmonisation and diagnostics of MIPAS ESA CH_4 and N_2O profiles using data assimilation, Atmos. Meas. Tech., 9, 5895–5909, https://doi.org/10.5194/amt-9-5895-2016, 2016.

Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia, A., Ehhalt, D., Endemann, M., Flaud, J. M., Gessner, R., Kleinert, A., Koopman, R., Langen, J., López-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., Perron, G., Remedios, J., Ridolfi, M., Stiller, G., and Zander, R.: MIPAS: an instrument for atmospheric and climate research, Atmos. Chem. Phys., 8, 2151–2188, https://doi.org/10.5194/acp-8-2151-2008, 2008.

Goldman, A., Paton-Walsh, C., Bell, W., Toon, G. C. andBlavier, J. F., Sen, B., Coffey, M. T., Hannigan, J. W., and Mankin, W. G.: Network for the detection of stratospheric change Fourier transform infrared intercomparison at Table Mountain Facility, November 1996, J. Geophys. Res., 104, 30 481–30 503, https://doi.org/10.1029/1999JD900879, 1999.

Hannigan, J. W., Coffey, M. T., and Goldman, A.: Semi-autonomous FTS observation system for stratospheric and tropospheric gases, J. Atmos. Ocean. Tech., 26, 1814–1828, https://doi.org/10.1175/2009JTECHA1230.1, 2009.

Hase, F., Hannigan, J. W., Coffey, M. T., Goldman, A., Hopfner, M., Jones, N. B., Rinsland, C. P., and Wood, S. W.: Intercomparison of retrieval codes used for the analysis of high-resolution, ground-based FTIR measurements, J. Quant. Spectrosc. Radiat. Transfer, 87, 25–52, https://doi.org/10.1016/j.jqsrt.2003.12.008, 2004.

Holl, G., Walker, K. A., Conway, S., Saitoh, N., Boone, C. D., Strong, K., and Drummond, J. R.: Methane cross-validation between three Fourier transform spectrometers: SCISAT ACE-FTS, GOSAT TANSO-FTS, and ground-based FTS measurements in the Canadian high Arctic, Atmos. Meas. Tech., 9, 1961–1980, https://doi.org/10.5194/amt-9-1961-2016, 2016.

Ishida, H. and Nakajima, T. Y.: Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager, J. Geophys. Res., 114, D07206, https://doi.org/10.1029/2008JD010710, 2009.

Ishida, H., Nakjima, T. Y., Yokota, T., Kikuchi, N., and Watanabe, H.: Investigation of GOSAT TANSO-CAI cloud screening ability through an intersatellite comparison, J. Appl. Meteorol. Climatol., 50, 1571–1586, https://doi.org/10.1175/2011JAMC2672.1, 2011.

Jin, J. J., Semeniuk, K., Beagley, S. R., Fomichev, V. I., Jonsson, A. I., McConnell, J. C., Urban, J., Murtagh, D., Manney, G. L., Boone, C. D., Bernath, P. F., Walker, K. A., Barret, B., Ricaud, P., and Dupuy, E.: Comparison of CMAM simulations of carbon monoxide (CO), nitrous oxide (N_2O), and methane (CH_4) with observations from Odin/SMR, ACE-FTS, and Aura/MLS, Atmos. Chem. Phys., 9, 3233–3252, 2009.

Kohlhepp, R., Ruhnke, R., Chipperfield, M. P., De Mazière, M., Notholt, J., Barthlott, S., Batchelor, R. L., Blatherwick, R. D., Blumenstock, T., Coffey, M. T., Demoulin, P., Fast, H., Feng, W., Goldman, A., Griffith, D. W. T., Hamann, K., Hannigan, J. W., Hase, F., Jones, N. B., Kagawa, A., Kaiser, I., Kasai, Y., Kirner, O., Kouker, W., Lindenmaier, R., Mahieu, E., Mittermeier, R. L., Monge-Sanz, B., Morino, I., Murata, I., Nakajima, H., Palm, M., Paton-Walsh, C., Raffalski, U., Reddmann, T., Rettinger, M., Rinsland, C. P., Rozanov, E., Schneider, M., Senten, C., Servais, C., Sinnhuber, B.-M., Smale, D., Strong, K., Sussmann, R., Taylor, J. R., Vanhaelewyn, G., Warneke, T., Whaley, C., Wiehle, M., and Wood, S. W.: Observed and simulated time evolution of HCl, $ClONO_2$, and HF total column abundances, Atmos. Chem. Phys., 12, 3527–3556, https://doi.org/10.5194/acp-12-3527-2012, 2012.

Kurylo, M. J. and Zander, R.: The NDSC–Its status after 10 years of operation, in: Proc. 19^{th} Quadrennial Ozone Symp., pp. 167–168, 2000.

Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, Appl. Opt., 48, 6716, https://doi.org/10.1364/AO.48.006716, 2009.

Kuze, A., Suto, H., Shiomi, K., Urabe, T., Nakajima, M., Yoshida, J., Kawashima, T., Yamamoto, Y., Kataoka, F., and Buijs, H.: Level 1 algorithms for TANSO on GOSAT: processing and on-orbit calibrations, Atmos. Meas. Tech., 5, 2447–2467, https://doi.org/10.5194/amt-5-2447-2012, 2012.

Kuze, A., Suto, H., Shiomi, K., Kawakami, S., Tanaka, M., Ueda, Y., Deguchi, A., Yoshida, J., Yamamoto, Y., Kataoka, F., Taylor, T. E., and Buijs, H. L.: Update on GOSAT TANSO-FTS performance, operations, and data products after more than 6 years in space, Atmos. Meas. Tech., 9, 2445–2461, https://doi.org/10.5194/amt-9-2445-2016, 2016.

Laeng, A., Plieninger, J., von Clarmann, T., Grabowski, U., Stiller, G., Eckert, E., Glatthor, N., Haenel, F., Kellmann, S., Kiefer, M., Linden, A., Lossow, S., Deaver, L., Engel, A., Hervig, M., Levin, I., McHugh, M., Noël, S., Toon, G., and Walker, K.: Validation of MIPAS IMK/IAA methane profiles, Atmos. Meas. Tech., 8, 5251–5261, https://doi.org/10.5194/amt-8-5251-2015, 2015.

Maksyutov, S., Patra, P. K., Onishi, R. a. S. T., and T., N.: NIES/FRCGC global atmospheric tracer transport model: Description, validation, and surface sources and sinks inversion, J. Earth Simul., 9, 3–18, 2008.

Notholt, J., Toon, G. C., Stordal, F., S., S., Schmidbauer, N., Becker, E., Meier, A., and Sen, B.: Seasonal variations of atmospheric trace gases in the high Arctic at 79° N, J. Geophys. Res., 102, 12 855–12 861, https://doi.org/10.1029/97JD00337, 1997.

Payan, S., Camy-Peyret, C., Oelhaf, H., Wetzel, G., Maucher, G., Keim, C., Pirre, M., Huret, N., Engel, A., Volk, M. C., Kuellmann, H., Kuttippurath, J., Cortesi, U., Bianchini, G., Mencaraglia, F., Raspollini, P., Redaelli, G., Vigouroux, C., de Mazière, M., Mikuteit, S., Blumenstock, T., Velazco, V., Notholt, J., Mahieu, E., Duchatelet, P., Smale, D., Wood, S., Jones, N., Piccolo, C., Payne, V., Bracher, A., Glatthor, N., Stiller, G., Grunow, K., Jeseck, P., Te, Y., and Butz, A.: Validation of version-4.61 methane and nitrous oxide observed by MIPAS, Atmos. Chem. Phys., 9, 413–442, https://doi.org/10.5194/acp-9-413-2009, 2009.

Picone, J. M., Hedin, A. E., Drob, D. P., and Aikin, A. C.: NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, J. Geophys. Res., 107, 1468, https://doi.org/10.1029/2002JA009430, 2002.

Plieninger, J., von Clarmann, T., Stiller, G. P., Grabowski, U., Glatthor, N., Kellmann, S., Linden, A., Haenel, F., Kiefer, M., Höpfner, M., Laeng, A., and Lossow, S.: Methane and nitrous oxide retrievals from MIPAS-ENVISAT, Atmos. Meas. Tech., 8, 4657–4670, https://doi.org/10.5194/amt-8-4657-2015, 2015.

Plieninger, J., Laeng, A., Lossow, S., von Clarmann, T., Stiller, G. P., Kellmann, S., Linden, A., Kiefer, M., Walker, K. A., Noël, S., Hervig, M. E., McHugh, M., Lambert, A., Urban, J., Elkins, J. W., and Murtagh, D.: Validation of revised methane and nitrous oxide profiles from MIPAS-ENVISAT, Atmos. Meas. Tech., 9, 765–779, https://doi.org/10.5194/amt-9-765-2016, 2016.

Pougatchev, N. S., Connor, B. J., and Rinsland, C. P.: Infrared measurements of the ozone vertical distribution above Kitt Peak, J. Geophys. Res., 100, 16, https://doi.org/10.1029/95JD01296, 1995.

Raspollini, P., Belotti, C., Burgess, A., Carli, B., Carlotti, M., Ceccherini, S., Dinelli, B. M., Dudhia, A., Flaud, J.-M., Funke, B., Höpfner, M., López-Puertas, M., Payne, V., Piccolo, C., Remedios, J. J., Ridolfi, M., and Spang, R.: MIPAS level 2 operational analysis, Atmos. Chem. Phys., 6, 5605–5630, https://doi.org/10.5194/acp-6-5605-2006, 2006.

Raspollini, P., Carli, B., Carlotti, M., Ceccherini, S., Dehn, A., Dinelli, B. M., Dudhia, A., Flaud, J.-M., López-Puertas, M., Niro, F., Remedios, J. J., Ridolfi, M., Sembhi, H., Sgheri, L., and von Clarmann, T.: Ten years of MIPAS measurements with ESA Level 2 processor V6 - Part 1: Retrieval algorithm and diagnostics of the products, Atmos. Meas. Tech., 6, 2419–2439, https://doi.org/10.5194/amt-6-2419-2013, 2013.

Raspollini, P., Arnone, E., Barbara, F., Carli, B., Castelli, E., Ceccherini, S., Dinelli, B. M., Dudhia, A., Kiefer, M., Papandrea, E., and Ridolfi, M.: Comparison of the MIPAS products obtained by four different level 2 processors, Ann. Geophys., 56, https://doi.org/10.4401/ag-6338, 2014.

Remedios, J. J., Leigh, R. J., Waterfall, A. M., Moore, D. P., Sembhi, H., Parkes, I., Greenhough, J., Chipperfield, M. P., and Hauglustaine, D.: MIPAS reference atmospheres and comparisons to V4.61/V4.62 MIPAS level 2 geophysical data sets, Atmos. Chem. Phys., 7, 9973–10 017, https://doi.org/10.5194/acpd-7-9973-2007, 2007.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, J. Geophys. Res., 108, 2156–2202, https://doi.org/10.1029/2002JD002299, 2003.

Rothman, L. S., Rinsland, C. P., Goldman, A., Massie, S. T., Edwards, D. P., Flaud, J.-M., Perrin, A., Camy-Peyret, C., Dana, V., Mandin, J.-Y., Schroeder, J., McCann, A., Gamache, R. R., Wattson, R. B., Yoshino, K., Chance, K., Jucks, K., Brown, L. R., Nemtchinov, V., and Varanasi, P.: The HITRAN Molecular Spectroscopic Database and HAWKS (HITRAN Atmospheric Workstation): 1996 Edition, J. Quant. Spectrosc. Radiat. Transfer, 60, 665–710, https://doi.org/10.1016/S0022-4073(98)00078-8, 1998.

Rothman, L. S., Jacquemart, D., Barbe, A., Chris Benner, D., Birk, M., Brown, L. R., Carleer, M. R., Chackerian, C., Chance, K., Coudert, L. H., Dana, V., Devi, V. M., Flaud, J.-M., Gamache, R. R., Goldman, A., Hartmann, J.-M., Jucks, K. W., Maki, A. G., Mandin, J.-Y., Massie, S. T., Orphal, J., Perrin, A., Rinsland, C. P., Smith, M. A. H., Tennyson, J., Tolchenov, R. N., Toth, R. A., Vander Auwera, J., Varanasi, P., and Wagner, G.: The HITRAN 2004 molecular spectroscopic database, J. Quant. Spectrosc. Radiat. Transfer, 96, 139–204, https://doi.org/10.1016/j.jqsrt.2004.10.008, 2005.

Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Boudon, V., Brown, L. R., Campargue, A., Champion, J.-P., Chance, K., Coudert, L. H., Dana, V., Devi, V. M., Fally, S., Flaud, J.-M., Gamache, R. R., Goldman, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W. J., Mandin, J.-Y., Massie, S. T., Mikhailenko, S. N., Miller, C. E., Moazzen-Ahmadi, N., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V. I., Perrin, A., Predoi-Cross, A., Rinsland, C. P., Rotger, M., Šimečková, M., Smith, M. A. H., Sung, K., Tashkun, S. A., Tennyson, J., Toth, R. A., Vandaele, A. C., and Vander Auwera, J.: The HITRAN 2008 molecular spectroscopic database, J. Quant. Spectrosc. Radiat. Transfer, 110, 533–572, https://doi.org/10.1016/j.jqsrt.2009.02.013, 2009.

Rothman, L. S., Gordon, I. E., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P. F., Birk, M., Bizzocchi, L., Boudon, V., Brown, L. R., Campargue, A., Chance, K., Cohen, E. A., Coudert, L. H., Devi, V. M., Drouin, B. J., Fayt, A., Flaud, J.-M., Gamache, R. R., Harrison, J. J., Hartmann, J.-M., Hill, C., Hodges, J. T., Jacquemart, D., Jolly, A., Lamouroux, J., Le Roy, R. J., Li, G., Long, D. A., Lyulin, O. M., Mackie, C. J., Massie, S. T., Mikhailenko, S., Müller, H. S. P., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V., Perrin, A., Polovtseva, E. R., Richard, C., Smith, M. A. H., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G. C., Tyuterev, V. G., and Wagner, G.: The HITRAN2012 molecular spectroscopic database, J. Quant. Spectrosc. Radiat. Transfer, 130, 4–50, https://doi.org/10.1016/j.jqsrt.2013.07.002, 2013.

Saeki, T., Saito, R., Belikov, D., and Maksyutov, S.: Global high-resolution simulations of CO_2 and CH_4 using a NIES transport model to produce a priori concentrations for use in satellite data retrievals, Geosci. Model Dev., 6, 81–100, https://doi.org/10.5194/gmd-6-81-2013, 2013.

Saitoh, N., Imasu, R., Ota, Y., and Niwa, Y.: CO_2 retrieval algorithm for the thermal infrared spectra of the Greenhouse Gases Observing Satellite: Potential of retrieving CO_2 vertical profile from high-resolution FTS sensor, J. Geophys. Res., 114, D17305, https://doi.org/10.1029/2008JD011500, 2009.

Saitoh, N., Kimoto, S., Sugimura, R., Imasu, R., Kawakami, S., Shiomi, K., Kuze, A., Machida, T., Sawa, Y., and Matsueda, H.: Algorithm update of the GOSAT/TANSO-FTS thermal infrared CO_2 product (version 1) and validation of the UTLS CO_2 data using CONTRAIL measurements, Atmos. Meas. Tech., 9, 2119–2134, https://doi.org/10.5194/amt-9-2119-2016, 2016.

Schneider, M., Blumenstock, T., Chipperfield, M., Hase, F., Kouker, W., Reddmann, T., Ruhnke, R., Cuevas, E., and Fischer, H.: Subtropical trace gas profiles determined by ground-based FTIR spectroscopy at Izaña (28° N, 16° W): Five year record, error analysis, and comparison with 3-D CTMs, Atmos. Chem. Phys., 5, 153–167, https://doi.org/10.5194/acp-5-153-2005, 2005.

Senten, C., de Mazière, M., Dils, B., Hermans, C., Kruglanski, M., Neefs, E., Scolas, F., Vandaele, A. C., Vanhaelewyn, G., Vigouroux, C., Carleer, M., Coheur, P. F., Fally, S., Barret, B., Baray, J. L., Delmas, R., Leveau, J., Metzger, J. M., Mahieu, E., Boone, C. D., Walker, K. A., Bernath, P. F., and Strong, K.: Technical Note: New ground-based FTIR measurements at Ile de La Réunion: observations, error analysis, and comparisons with independent data, Atmos. Chem. Phys., 8, 3483–3508, https://doi.org/10.5194/acp-8-3483-2008, 2008.

Sepúlveda, E., Schneider, M., Hase, F., García, O. E., Gomez-Pelaez, A., Dohe, S., Blumenstock, T., and Guerra, J. C.: Long-term validation of tropospheric column-averaged CH_4 mole fractions obtained by mid-infrared ground-based FTIR spectrometry, Atmos. Meas. Tech., 5, 1425–1441, https://doi.org/10.5194/amt-5-1425-2012, 2012.

Sepúlveda, E., Schneider, M., Hase, F., Barthlott, S., Dubravica, D., García, O. E., Gomez-Pelaez, A., González, Y., Guerra, J. C., Gisi, M., Kohlhepp, R., Dohe, S., Blumenstock, T., Strong, K., Weaver, D., Palm, M., Sadeghi, A., Deutscher, N. M., Warneke, T., Notholt, J., Jones, N., Griffith, D. W. T., Smale, D., Brailsford, G. W., Robinson, J., Meinhardt, F., Steinbacher, M., Aalto, T., and Worthy, D.: Tropospheric CH_4 signals as observed by NDACC FTIR at globally distributed sites and comparison to GAW surface in situ measurements, Atmos. Meas. Tech., 7, 2337–2360, https://doi.org/10.5194/amt-7-2337-2014, 2014.

Sheese, P., Walker, K., and Boone, C.: Atmospheric pseudo-retrievals for averaging kernel and total uncertainty characterization for ACE-FTS level 2 (PRAKTICAL) data, in: EGU General Assembly, vol. 18 of *EGU General Assembly*, p. 17582, 2016.

Sheese, P. E., Boone, C. D., and Walker, K. A.: Detecting physically unrealistic outliers in ACE-FTS atmospheric measurements, Atmos. Meas. Tech., 8, 741–750, https://doi.org/10.5194/amt-8-741-2015, 2015.

Sussmann, R. and Schäfer, K.: Infrared spectroscopy of tropospheric trace gases: combined analysis of horizontal and vertical column abundances, Appl. Opt., 36, 735–741, https://doi.org/10.1364/AO.36.000735, 1997.

Sussmann, R., Forster, F., Rettinger, M., and Jones, N.: Strategy for high-accuracy-and-precision retrieval of atmospheric methane from the mid-infrared FTIR network, Atmos. Meas. Tech., 4, 1943–1964, https://doi.org/10.5194/amt-4-1943-2011, 2011.

Sussmann, R., Ostler, A., Forster, F., Rettinger, M., Deutscher, N. M., Griffith, D. W. T., Hannigan, J. W., Jones, N., and Patra, P. K.: First intercalibration of column-averaged methane from the Total Carbon Column Observing Network and the Network for the Detection of Atmospheric Composition Change, Atmos. Meas. Tech., 6, 397–418, https://doi.org/10.5194/amt-6-397-2013, 2013.

Toon, G. C., Blavier, J.-F., Sung, K., Rothman, L. S., and E. Gordon, I.: HITRAN spectroscopy evaluation using solar occultation FTIR spectra, J. Quant. Spectrosc. Radiat. Transfer, 182, 324–336, https://doi.org/10.1016/j.jqsrt.2016.05.021, 2016.

Vincenty, T.: Direct and Inverse Solutions of Geodesics on the Ellipsoid with application of nested equations, Survey Review, 23, 88–93, https://doi.org/10.1179/sre.1975.23.176.88, 1975.

von Clarmann, T., Höpfner, M., Kellmann, S., Linden, A., Chauhan, S., Funke, B., Grabowski, U., Glatthor, N., Kiefer, M., Schieferdecker, T., Stiller, G. P., and Versick, S.: Retrieval of temperature, H_2O, O_3, HNO_3, CH_4, N_2O, $ClONO_2$ and ClO from MIPAS reduced resolution nominal mode limb emission measurements, Atmos. Meas. Tech., 2, 159–175, https://doi.org/10.5194/amt-2-159-2009, 2009.

Waymark, C., Walker, K. A., Boone, C. D., and Bernath, P. F.: ACE-FTS version 3.0 data set: validation and data processing update, Ann. Geophys., 56, https://doi.org/10.4401/ag-6339, 2013.

Wiacek, A., Taylor, J. R., Strong, K., Saari, R., Kerzenmacher, T., Jones, N. B., and Griffith, D. W. T.: Ground-Based solar absorption FTIR spectroscopy: characterization of retrievals and first results from a novel optical design instrument at a New NDACC Complementary Station, J. Atmos. Ocean. Tech., 24, 432–448, https://doi.org/10.1175/JTECH1962.1, 2007.

Wood, S. W., Bodeker, G. E., Boyd, I. S., Jones, N. B., Connor, B. J., Johnston, P. V., Matthews, W. A., Nichol, S. E., Murcray, F. J., Nakajima, H., and Sasano, Y.: Validation of version 5.20 ILAS HNO_3, CH_4, N_2O, O_3, and NO_2 using ground-based measurements at Arrival Heights and Kiruna, J. Geophys. Res., 107, (D24)8208, https://doi.org/10.1029/2001JD000581, 2002.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global Concentrations of CO_2 and CH_4 Retrieved from GOSAT: First Preliminary Results, Sci. Online Lett. Atmos., 5, 160–163, https://doi.org/10.2151/sola.2009-041, 2009.

Zander, R., Mahieu, E., Demoulin, P., Duchatelet, P., Roland, G., Servais, C., de Mazière, M., Reimann, S., and Rinsland, C. P.: Our changing atmosphere: evidence based on long-term infrared solar observations at the Jungfraujoch since 1950, Sci. Total Environ., 391, 184–195, https://doi.org/10.1016/j.scitotenv.2007.10.018, 2008.

Zou, M., Xiong, X., Saitoh, N., Warner, J., Zhang, Y., Chen, L., Weng, F., and Fan, M.: Satellite observation of atmospheric methane: intercomparison between AIRS and GOSAT TANSO-FTS retrievals, Atmos. Meas. Tech., 9, 3567–3576, https://doi.org/10.5194/amt-9-3567-2016, 2016.

Table 1. FTS instruments used in the CH_4 VMR vertical profile comparisons presented herein.

Instrument	Spectral Resolution[a]	Spectral Range[b]	Viewing Geometry	NDACC Latitude	NDACC Longitude	Reference
TANSO-FTS	0.2 cm^{-1}	700–1800 cm^{-1}	nadir			Kuze et al. (2009)
MIPAS	0.0625 cm^{-1}	685–2410[c] cm^{-1}	limb			Fischer et al. (2008)
ACE-FTS	0.02 cm^{-1}	750–4400 cm^{-1}	solar occultation			Bernath et al. (2005)
Eureka	0.0024 cm^{-1}	450–4800 cm^{-1}	ground	80.1° N	86.4° W	Batchelor et al. (2009)
Ny Ålesund	0.0015 cm^{-1}	475–4500 cm^{-1}	ground	78.9° N	11.9° E	Notholt et al. (1997)
Thule	0.004 cm^{-1}	700–5000 cm^{-1}	ground	76.5° N	68.8° W	Goldman et al. (1999)
Kiruna	0.0024 cm^{-1}	450–4800 cm^{-1}	ground	67.8° N	20.4° E	Blumenstock et al. (2006)
Bremen	0.0024 cm^{-1}	450–4800 cm^{-1}	ground	53.1° N	8.8° E	Buchwitz et al. (2007)
Zugspitze	0.0015 cm^{-1}	475–4500 cm^{-1}	ground	47.4° N	11.0° E	Sussmann and Schäfer (1997)
Jungfraujoch	0.0015 cm^{-1}	475–4500 cm^{-1}	ground	46.6° N	8.0° E	Zander et al. (2008)
Toronto	0.004 cm^{-1}	750–8500 cm^{-1}	ground	43.6° N	79.4° W	Wiacek et al. (2007)
Izaña	0.0024 cm^{-1}	450–4800 cm^{-1}	ground	28.3° N	16.5° W	Schneider et al. (2005)
Mauna Loa	0.0024 cm^{-1}	450–4800 cm^{-1}	ground	19.5° N	155.6° W	Hannigan et al. (2009)
Altzomoni[d]	0.0024 cm^{-1}	450–4800 cm^{-1}	ground	19.1° N	98.7° W	Baylon et al. (2014)
St. Denis, La Réunion	0.0036 cm^{-1}	600–4300 cm^{-1}	ground	20.9° S	55.5° E	Senten et al. (2008)
Maïdo, La Réunion[e]	0.0024 cm^{-1}	600–4500 cm^{-1}	ground	21.1° S	55.4° E	Baray et al. (2013)
Wollongong	0.0024 cm^{-1}	450–4800 cm^{-1}	ground	34.4° S	150.9° E	Kohlhepp et al. (2012)
Lauder	0.0035 cm^{-1}	700–4500 cm^{-1}	ground	45.0° S	169.7° E	Bader et al. (2017)
Arrival Heights	0.0035 cm^{-1}	750–4500 cm^{-1}	ground	77.8° S	166.6° E	Wood et al. (2002)

[a] For NDACC instruments, the best achievable spectral resolution is listed here. Operationally achieved spectral resolutions for NDACC instruments may be coarser.

[b] NDACC instruments use optical filters that reduce the effective spectral range when making measurements.

[c] MIPAS' spectral resolution is divided into four, narrower bands.

[d] The Altzomoni site came online in late 2012.

[e] The Maïdo, La Réunion site came online in early 2013.

Table 2. Number of coincident CH_4 VMR vertical profile measurements that were found between TANSO-FTS retrievals and those from ESA MIPAS, IMK-IAA MIPAS, ACE-FTS, and NDACC stations. The three columns show the total number of coincidences found, the number of unique TANSO-FTS measurements within those coincidences, and the size of the reduced, one-to-one coincidences used.

Target Instrument	Total Coincident Profiles	Unique TANSO-FTS Profiles	One-to-one Profiles Used
ESA MIPAS	450,230	358,267	85,386
IMK-IAA MIPAS	267,065	210,573	51,099
ACE-FTS	51,937	47,560	4,302
Total NDACC	213,181	44,920	17,637
Eureka	11,843	2,447	1,009
Ny Ålesund	5,445	1,300	349
Thule	6,997	3,359	513
Kiruna	4,595	2,056	529
Bremen	2,610	1,452	211
Zugspitze	47,512	5,743	3,469
Jungfraujoch	18,757	5,938	1,493
Toronto	9,909	5,195	816
Izaña	56,254	4,336	4,501[a]
Mauna Loa	4,338	2,381	379
Altzomoni	4,746	854	486
St. Denis, La Réunion	12,270	3,161	1,507
Maïdo, La Réunion	3,139	868	383
Wollongong	27,781	4,808	2,365
Lauder	7,083	2,638	704
Arrival Heights	5,042	3,122	258

[a] The Izaña NDACC coincidence data set is the only one in which TANSO-FTS measurements are more sparse. For consistency, Izaña was not treated as a special case.

Table 3. Statistics for the partial column integration ranges for ESA MIPAS, IMK-IAA MIPAS, ACE-FTS and NDACC stations with the requirements that the TANSO-FTS sensitivity, s, is greater than 0.2 for at least three pressure levels. The number of coincident profiles passing this criterion, N, and its percentage of one-to-one coincidences found in this study are given. Means and standard deviations are given for the minimum altitudes, $\min(z)$, total integration range, z_{range}, and number of levels used, n.

Target	Profiles with $s > 0.2$		Lowest Altitude (km)		Altitude Range (km)		Number of Levels	
Instrument	N	(%)	$\min(z)$	$\sigma_{\min(z)}$	z_{range}	$\sigma_{z_{range}}$	n	σ_n
ESA MIPAS	52,016	60.9	8.4	1.5	4.6	1.5	4.8	1.1
IMK-IAA MIPAS	17,787	34.8	11.3	0.6	3.5	0.9	3.7	0.6
ACE-FTS	2,562	59.6	7.3	1.4	5.2	2.3	5.4	1.8
Total NDACC	18,587	98.0	3.3	1.0	11.3	2.1	10.4	1.5

Table 4. Least squares regression statistics for the data in each hemisphere plotted in Fig. 10. Results from all four validation target ~~datasets~~ data sets are combined.

	Slope (~~ppmv~~ppbv/° latitude)	Intercept (~~ppmv~~ppbv)	R^2
Northern	~~0.000113 ± 0.000005~~ 0.113 ± 0.005	~~0.0053 ± 0.0003~~ 5.3 ± 0.3	0.08
Southern	~~-0.000207 ± 0.000004~~ -0.207 ± 0.004	~~0.0031 ± 0.0002~~ 3.1 ± 0.2	0.18

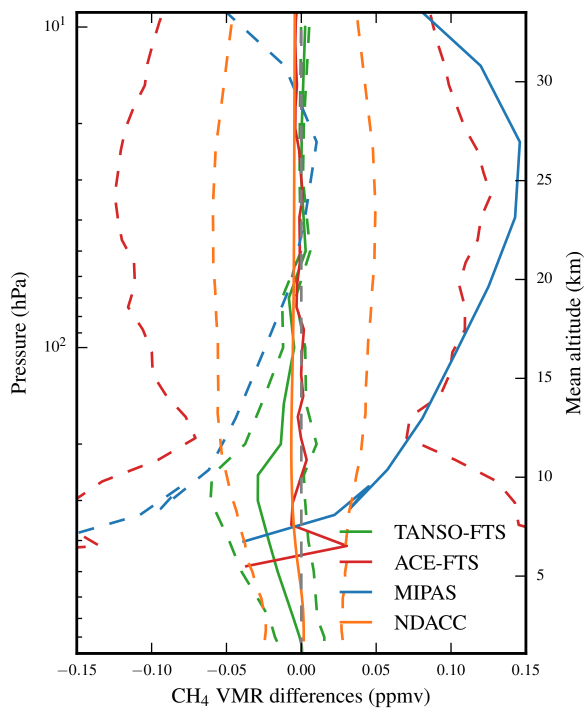

Figure 1. Results for investigating the ~~internal~~ variability ~~of~~ within each CH_4 VMR profile data set~~, comparing~~. Shown are the following comparisons: TANSO-FTS ~~measurements with~~ retrievals compared to their a priori (green), pairs of sequential ACE-FTS ~~measurements~~ retrievals (red), ESA MIPAS ~~measurements with~~ retrievals to IMK-IAA ~~measurements~~ MIPAS retrievals made for the same limb observations (blue), and pairs of NDACC ~~measurements with others~~ retrievals made on the same day (orange). All retrieved profiles used are coincident with TANSO-FTS. Dashed lines are one standard deviation.

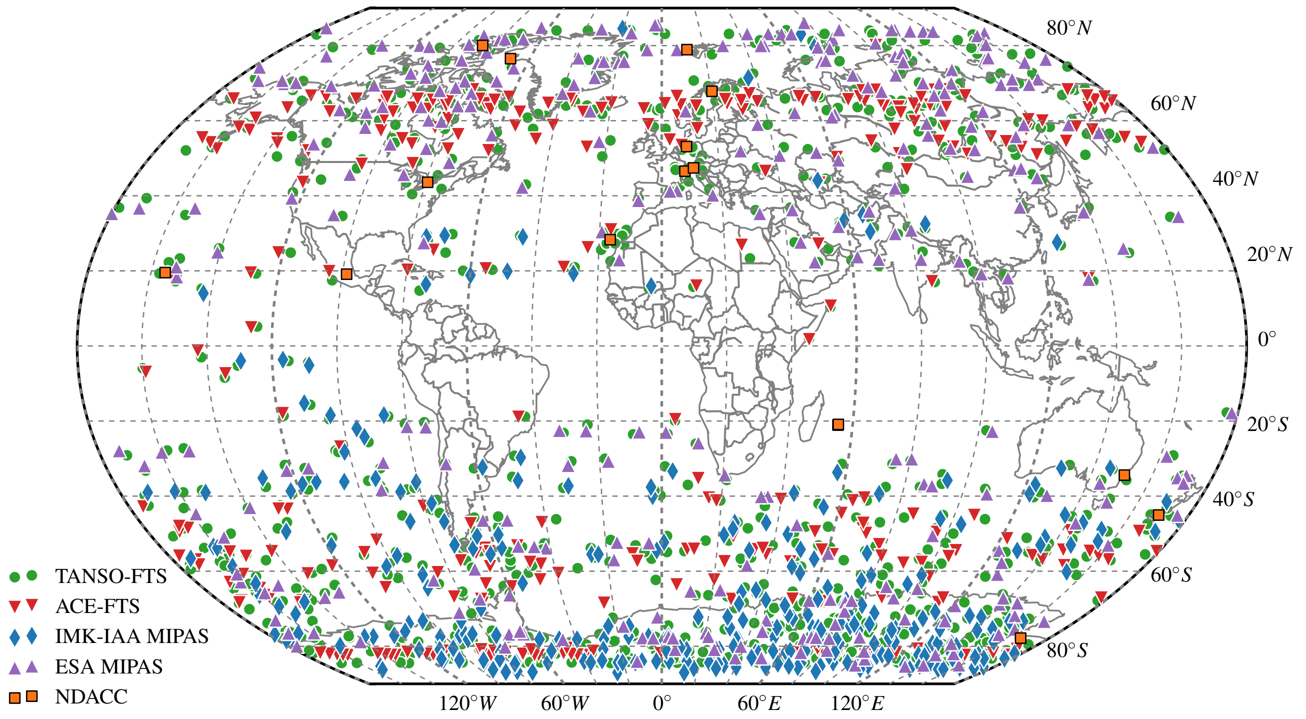

Figure 2. Locations of the first 200 observations of 2012 used in this study for TANSO-FTS (green), ACE-FTS (red), and IMK-IAA MIPAS (blue), ESA MIPAS (purple). The NDACC stations are shown in orange.

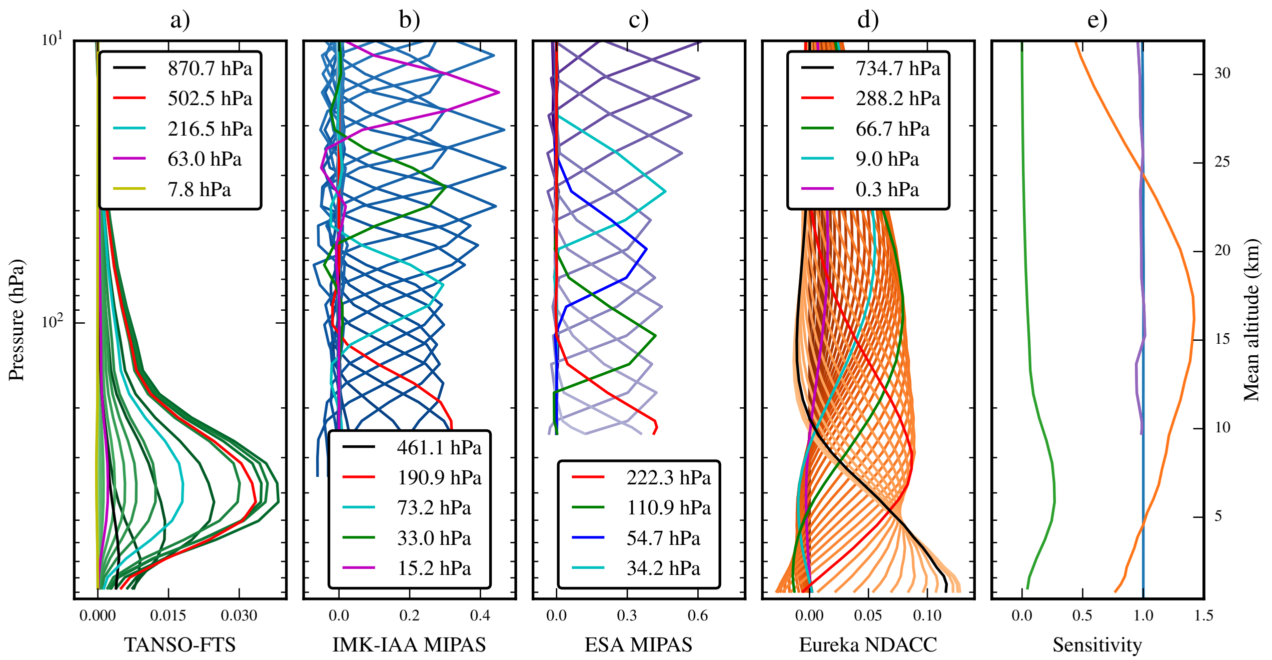

Figure 3. Example of averaging kernels for: a) TANSO-FTS, b) IMK-IAA MIPAS, c) ESA MIPAS, and d) NDACC. Each kernel shown is the mean from 30 averaging kernel matrices from measurements made over the Arctic, interpolated to a common pressure grid. Panel d) shows the mean averaging kernels from the Eureka station. Panel e) shows the sensitivity for the mean averaging kernels shown in each panel: TANSO-FTS (green), IMK-IAA MIPAS (blue), ESA MIPAS (purple), and NDACC (orange).

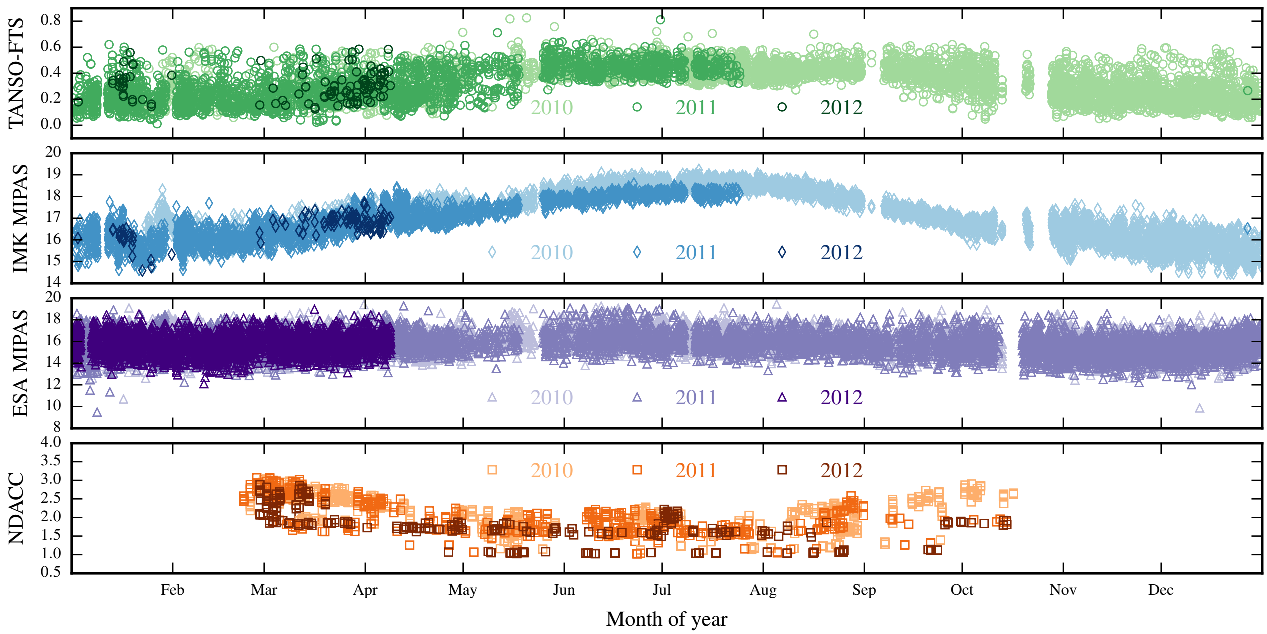

Figure 4. Degrees of freedom for signal for, from top to bottom: TANSO-FTS, IMK-IAA MIPAS, ESA MIPAS, and NDACC. Each satellite (and panel) uses a different symbol and colour, but the colour shades indicate the year the measurement was made in. The TANSO-FTS and IMK-IAA MIPAS measurements shown are in coincidence. The ESA MIPAS and NDACC data are from our analyzed data set, but not in coincidence with the TANSO-FTS data in the top panel. All data are from the Arctic, 90–60° N, with the NDACC measurements from Eureka, Ny Ålesund and Thule.

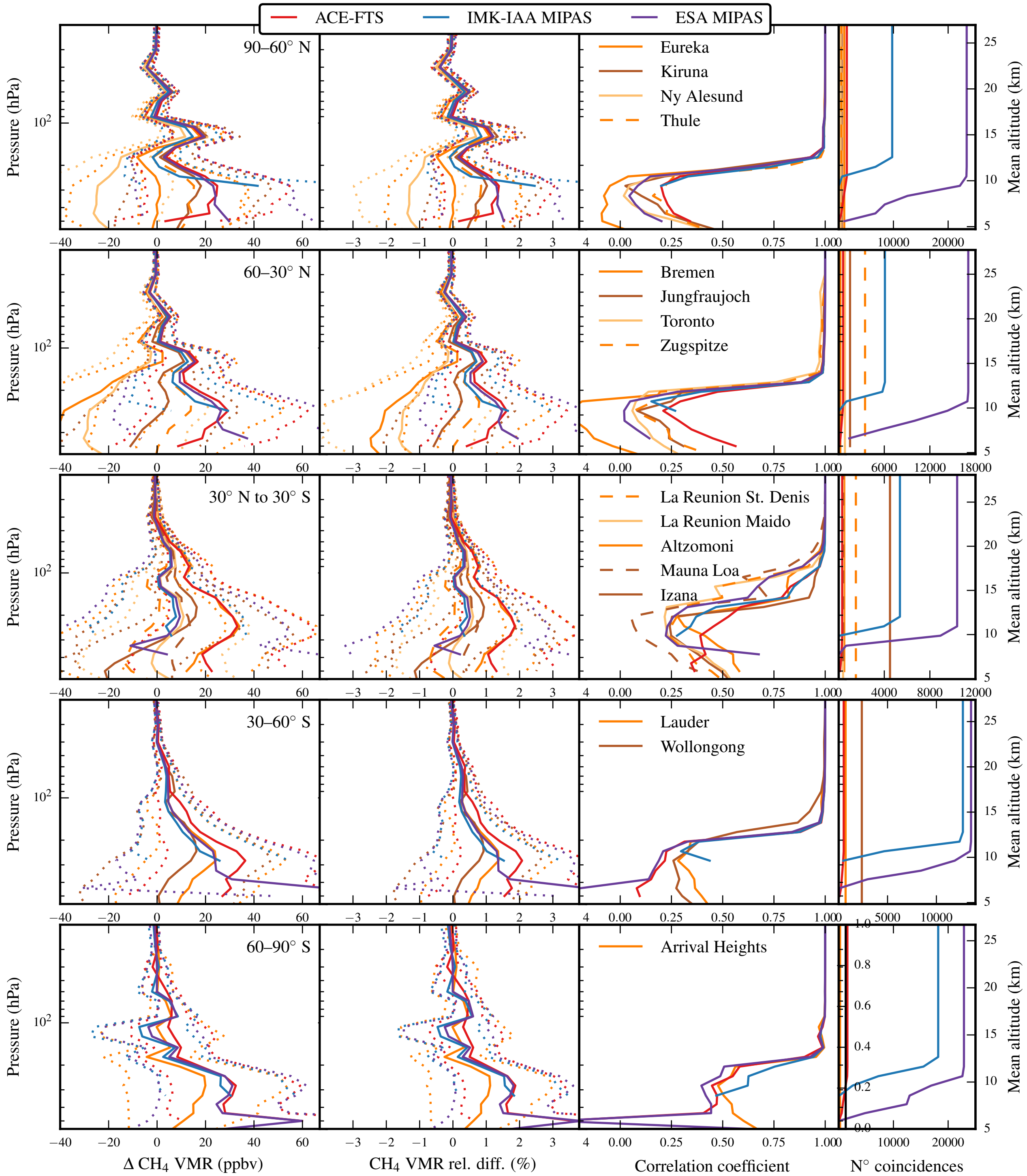

Figure 5. Zonally averaged comparison results. The rows present results for each zone, from top to bottom: 90–60° N, 60–30° N, 30° N–30° S, 30–60° S, and 60–90° S. In each row, the four panels show, from left to right, the mean CH_4 VMR difference between retrievals from TANSO-FTS and the validation target at each pressure level; the mean CH_4 VMR differences relative to the mean CH_4 VMR vertical profile of the validation target; the correlation coefficients R^2 of the CH_4 VMR differences for each coincident pair at each pressure level; and the number of coincidences at each pressure level. Differences are calculated as TANSO-FTS $-$ $target$ for each ~~dataset~~ data set compared. In

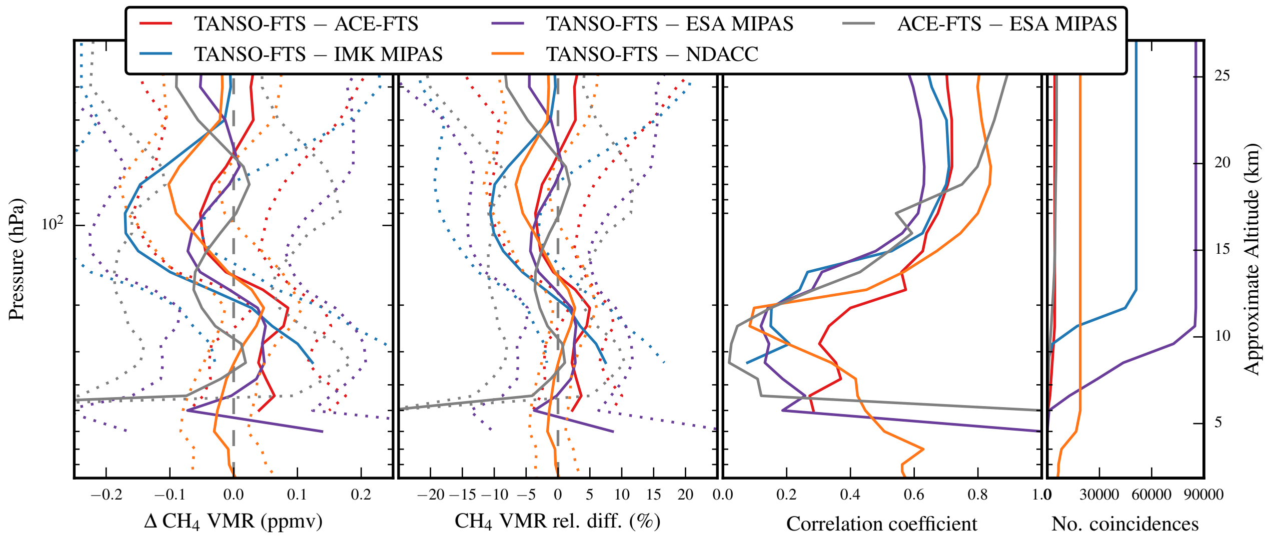

Figure 6. Averaged comparison results, as in each panel of Fig. 5, for all latitudes, without applying smoothing to the validation instruments' CH_4 VMR vertical profiles. Differences are calculated as TANSO-FTS − $target$ for each ~~dataset~~ data set compared (and ACE-FTS − ESA MIPAS for that case).

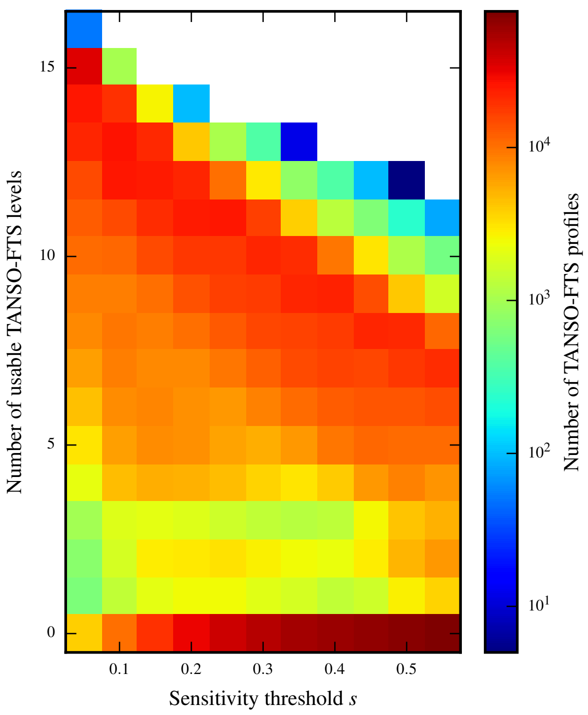

Figure 7. Two-dimensional histogram showing the number of TANSO-FTS CH_4 VMR profiles within our data set (z-axis) that have some number of usable pressure levels (y-axis) with a sensitivity greater than some given threshold, s (x-axis). The data set shown here consists of all TANSO-FTS observations that are one-to-one coincident with a target validation ~~dataset~~data set. The threshold chosen for this study was $s = 0.2$.

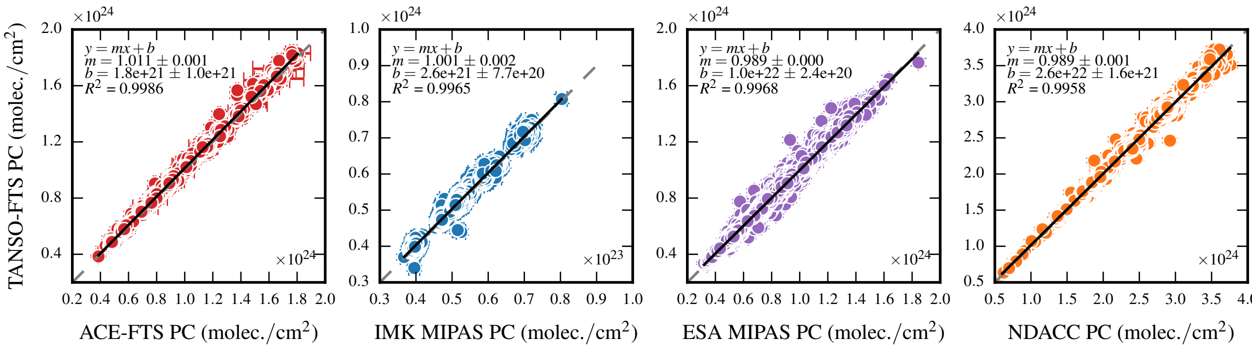

Figure 8. Partial column (PC) correlation plots comparing TANSO-FTS CH_4 to each validation instrument. Comparisons to ACE-FTS are red, to IMK-IAA MIPAS are blue, to ESA MIPAS are purple, and to NDACC are orange. The vertical range of partial column integration varies for each pair of coincident profiles based on the criteria described in Sect.7.1. The statistics for weighted linear least-squares regression are shown, with weights equal to $1/(\delta_x^2 + \delta_y^2)$.

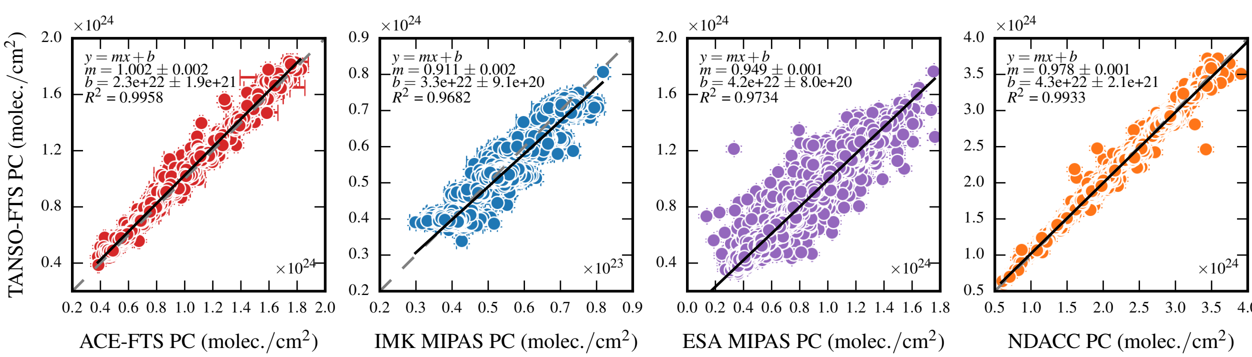

Figure 9. As in Fig. 8, but for partial column correlation results using unsmoothed CH_4 VMR vertical profiles for each validation instrument.

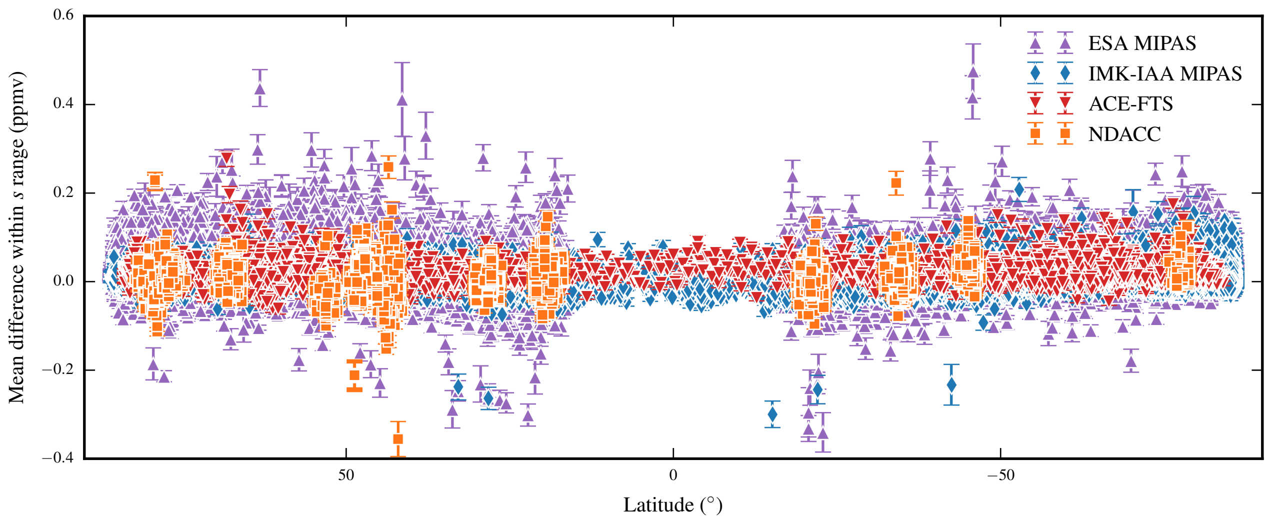

Figure 10. Mean CH_4 VMR differences between TANSO-FTS and each validation target ~~dataset~~data set, averaged vertically using the altitude range selected for integrating partial columns as a function of latitude. Differences are calculated as TANSO-FTS $-$ $target$ for each ~~dataset~~ data set compared.