Dear Reviewer,

Thank you for your comprehensive review of our paper, below is our response to your review. We have kept your original comments in black, our responses are in blue, and specific actions are indicated in underlined blue.

The work entitled "On the consistency of methane retrievals using TCCON and multiple spectroscopic databases" by Edward Molina et al. treats a highly relevant topic and suits well within the scope of AMT. While the empirical database used (time series of high-resolution FTIR spectra collected at four TCCON sites in different latitudes covering the year 2020) is a sound observational basis for the conducted evaluation of different spectroscopic databases and the overall structure of the paper is appropriate, there are significant shortcomings in the methodological concepts. The manuscript requires major revisions.

Main points of criticism, which need consideration, are:
The presented discussion based on methane dry air molar fractions (DMFs) ignores the fact that this product is based on the ratio of the target gas and molecular oxygen. In order to compensate for spectroscopic uncertainties in both methane and oxygen, the reported DMFs involves empirical airmass independent and airmass-dependent calibrations.
Therefore, when comparing the accurateness of methane band intensities or evaluating spurious air-mass dependencies triggered by imperfect description of pressure-broadening effects in different line lists, the effects introduced by the oxygen (which brings in both a considerable calibration bias and an airmass dependency) and the empirical post-corrections of the GGG processing chain need to be avoided. I believe that for the desired purpose molar fractions should instead be constructed from the column amount of the target gas, the water vapour column and the recorded ground pressure.

The empirical post-correction incorporated in GGG2014 is made to harmonize with the associated oxygen and methane line lists. It is therefore a trivial finding that the GGG2014 line list shows small airmass dependence and that other line lists create a calibration offset. I would expect from a study of the kind presented to demonstrate which methane line list has the best performance (correct calibration of band strengths and minimal SZA dependence of DMF). While the operational GGG2014 results (using the foreseen line lists and the corrections in the post-processing chain) present the best available reference for the true DMF value, the DMF values for the investigation of different line lists and micro windows should be constructed as described above. Note that using this procedure the GGG2014 methane line list can be tested in an equivalent manner to all other line lists versus the operational DMF as target quantity. This would reveal the actual calibration biases of the methane band intensities and identify the methane line list with the most realistic description of pressure broadening effects (which would create the least SZA dependence).

Thank you for this important point. We agree with the reviewer here that using the DMF values do introduce biases from the O2 spectroscopic parameters, and the empirical post-corrections. However, we disagree with the approach of using the surface pressure and water vapour columns to calculate the VMRs. One of the key benefits of using the O2 column to calculate DMFs is that it is a constant in the atmosphere, meaning that we can directly compare the DMFs calculated from each site. If we calculate VMRs for each site, these are no longer directly comparable, since we introduce biases from the surface pressure measurement, and from the water column calculation (used for calculating the dry air). This means our comparisons would then become dependent on the uncertainties associated with water vapour and surface pressure from each site, which will be larger than O2. Further, the empirical post-correction is applied to all of the retrievals, meaning the applied bias will always be the same, rather than varying as would be the case with VMR calculations.

Fundamentally, although we agree with the point that using O2 introduces biases into our results, we think using VMR would add additional uncertainties that would make the results less valuable. We have now introduced this argument as a point in the discussion of the paper, and we propose an additional assessment using VMRs be caried out.

The other major shortcoming of the work is that the line lists are not investigated species-wise. The study would be of much higher value and more conclusive, if the water vapour line list would be investigated independently from the methane line list. I understand this would require a number of additional retrieval runs, but would be worth the effort as the work then would provide real guidance for selecting both the best methane and water vapour line lists. In its current form, it is difficult to see whether a temperature or humidity dependence is generated by shortcomings of the methane line list itself or whether it is just a consequence of a poor description of interfering water vapour lines (exceedingly important in window 1). I would strongly recommend to include this kind of extension and instead skip the included shallow discussion of methane isotopologue retrievals, which results in the implausible suggestion to use GEISA2020 for this purpose, although this line list on the whole does not appear to be a great choice even for the main isotopologue.

We agree with the reviewer that adding a similar investigation of water vapour (and probably HDO as a separate species) would greatly enhance this study, and is certainly a necessary next step. However, we think this would require an additional paper to describe this work, since this water vapour study would require an assessment as in-depth as that already shown in this paper. In addition to assessing the interfering water vapour lines in the methane windows, this study should also include an assessment of the water vapour windows nominally used in TCCON, which are significantly different to the methane windows. Therefore, since this paper is already of significant length and depth, we feel this study would yield a paper that tries to cover too much ground.

Detailed comments:
In several figures, the data seem to be clipped in a significant manner, so information might be lost (spectral residual in figure 1, bar charts in figures 2 and 7 (impossible to decide whether the bars end or are clipped), time series Figure 5 (I agree to use the same scale for all figures, but too many Tsukuba data points are lost with the current choice)).

Thanks you for this point, we have now modified the y-axes in all of these figures so that no data is clipped.

The implementation of the GGG2020 spectroscopic data is not described properly: are SDV and line-mixing parameters taken into account? While the authors state in section they use the "GGG2014 environment" and state in section 4 that the GGG2020 software uses "non-Voigt line shapes for methane" and could therefore yield "improved or different results", they on the other hand state in section 2.3 that they implemented the qSDHC model and first order line mixing. So I do not see why GGG2020 would generate other results out of the same line list. A proper comparison of the GGG2020 line list versus other lists in the modified GGG2014 environment certainly needs to include the non-Voigt extensions. The description needs to be clarified.

Thank you for highlighting this point of confusion, while GGG2020 does include non-Voigt parameters, the GGG2014 software used in this study was not modified to take advantage of them, the software was only modified to take advantage of the parameters in the SEOM-IAS database. We have now modified the relevant point in section 2.3 and the discussion to highlight the fact that these parameters are not used in this study, and may provide improved results if implemented in the future.

While figure 2 suggests that SEOM-IAS outperforms the other line lists in window 1, this is not supported by figure 1. I understand that figure 1 is just an example, but it would be better to show a typical example (it might be instructive to include in the appendix an equivalent figure for each site).

We agree, we have changed the example plot, and we have identified in Figure 2 (which has changed a little bit to reflect a wider spectra sample) that SEOM-IAS and GGG2020 show more or less equivalent results. We have also included typical spectra from each site in the appendix.

The chosen performance indicators for the spectral fit quality introduce dependence on the SNR of the measured spectra (this propagates also into the DMF indicators due to the variable retrieval error, if I understand correctly). Time series of SNR or at least typical SNR values achieved at each site should be provided. The SNR would be determined from transparent spectral sections in the vicinity of the micro windows (two values might be sufficient, but the SNR in the window 1 region might differ significantly from the other windows).

Thank you for this point, SNR was typically not calculated with GGG2014 as there was no consensus in the TCCON community as how to exactly calculate it for the TCCON retrievals, although we understand that GGG2020 will provide this opportunity. Unfortunately, the output from GGG2014 does not include any of the transparent spectral regions we would need to calculate SNR, and only show the micro-windows applied in the retrieval process. The raw spectra input into GGG2014 are in a binary format, and are there unable to be easily read and understood.

Two aspects of the shown TCCON methane time series seem suspicious: HITRAN2016, GEISA2020, and SEOM-IAS show a high bias with respect to GGG2014 at most sites. Surprisingly, this bias is absent in the Ascension time series. Is there an explanation for this behaviour? The Darwin XCH4 time series shows an uncommented 1% (!) step change to higher values from March onwards. Is this a geophysical signal or a problem with data quality? In addition to methane, the time series of XAIR as a quality indicator should be added as additional panel to the figure for all sites.

There are a number of points to discuss in this comment which we answer in the following numbered items.

1) W.r.t the lack of a bias in the Ascension Island results, we think there are a number of reasons for this. Firstly it is possible that the Ascension Island instrument is less noisy than the other TCCON instruments. Secondly and more likely, the Ascension Island data used in this study unlike the other sites was hand picked for 'golden days' only, so only high quality retrievals were used. In addition, the conditions at Ascension Island hardly vary, while the other sites experience more challenging and varying conditions, which could explain the differences.

2) The Darwin results are a little misleading, we studied this 1% step change carefully and only the first few weeks of March appear to be affected by this change. Which is apparent by the relative lack of variability in the retrieval results when compared to the remainder of the time series. All typical quality control indicators (XAir etc) do not flag this period as being of poor quality, and it is not easily apparent as to why this abnormality exists. Because none of the standard indicators do not suggest anything wrong with these retrievals, we have not filtered them from the results, but <u>we have highlighted it in our updated paper</u>.

3) <u>We have now added XAir as an additional panel for each site as an indicator of the quality.</u> We show XAir for GGG2014, 2020, HITRAN and GEISA2020 as SEOM-IAS does not have any spectral lines in the TCCON XAir retrieval window.

While a linear fit model is certainly appropriate for quantifying the impact of temperature, this probably is much less so for the airmass dependency. It would be good to add figures for each window and line list showing the typical SZA (or airmass) dependency (for clearly showing the effect, the daily data of a station time series would be normalized to a daily reference value at a chosen intermediate SZA typically covered by all sites, then the ratio would be plotted for the complete time series). Such figures would reveal the strength of the effect and inspire an appropriate functional form for fitting the dependency.

Thank you for this important point, we agree that SZA dependency is likely not a linear relationship. In order to further investigate possible fit models we plotted the histograms SZA, using the normalized residual between the retrieved window/spectroscopic database as weights. This helped identified the relative sensitivity of biases to SZA, which we generally found to be more of a second order relationship rather than a first order one. We decided not to include the histograms into the paper, since we feel this would add numerous figures to an already lengthy paper. We have replaced the original figure 8, with the updated fit models, which clearly indicate non-linear relationships for SZA. Out of curiosity we applied second order fits to the other quantities, and found that in some cases temperature exhibited non-linear relationships. We have therefore included these fits into the updated figure 8 (now figure 9) into the paper as well.

Dear Reviewer,

Thank you for your comprehensive review of our paper, below is our response to your review. We have kept your original comments in black, our responses are in blue, and specific actions are indicated in underlined blue.

This paper details the impact of using 5 different spectroscopic databases (GGG2014, GGG2020, GEISA2020, HITRAN2016, SEOM-IAS) on the retrieval of 12CH4 DMFs in 4 different spectral windows from TCCON measurements. The retrieval is performed with the GFIT retrieval algorithm from the GGG2014 environment. Retrievals are carried out on approximately a year of data at 4 TCCON stations (Ascension Island, Darwin, Ny Alesund, Tsukuba), chosen to have an optimal coverage of different atmospheric and measurement conditions (temperature, water vapour, SZA). A second objective of the paper is also to assess the impact of the spectroscopy and fit windows on the retrieval of 13CH4 (and indirectly, the d13C) and check for consistency of the results.

The scientific work is extensive, with a lot of data processing performed to represent retrievals under various atmospheric and measuring conditions. The presentation and analysis of the data is sometimes a little laborious, since so many different parameters are considered (spectroscopic database, retrieval window, TCCON sites). Still, the paper mostly presents the results adequately and underlines the important results. The results are scientifically relevant and interesting, and are certainly within the scope of AMT.

However, the paper sometimes lacks in style and clarity. It is obvious that it has been through some re-writes and major changes, and it should be proof-read and improved in terms of flow and clarity. I would recommend publication of the paper in AMT pending some minor revisions of the manuscript, as detailed below. Note that I started to note down many typos or problematic sentences in the first half of the paper but was not as systematic in the second half and would recommend to proof-read this more thoroughly.

Thank you for your comments, we have responded to your criticisms in line below.

Comments

1/ The abstract should be improved.

a) The third paragraph is especially difficult to read. It would be helpful to be more quantitative in your description of the impact. The sentence “We also find strong evidence that different windows in different spectroscopic databases exhibit different levels of sensitivity to changing local conditions such as light path length and water vapour” is technically correct, but it is so general it becomes almost meaningless.

b) You should mention the work on the 13CH4 retrieval and d13C determination.

Thank you for this point, we have re-written the abstract to account for your comments a) and b) here.

2/ The introduction should be re-worked to improve clarity and fluidity:

a) I find it strange that a paper studying the impact of the spectroscopy on the retrieval of methane does not directly discuss the spectroscopy aspect of the retrieval in its introduction. Maybe add 1-2 lines to describe how it is relevant?

Thank you for pointing this out, we have now added the following text in the first paragraph to highlight the spectroscopic aspect.

The remote sensing of methane is fundamentally dependent on inferring the concentrations from the absorption of light in the atmosphere at wavelengths unique to methane, otherwise known as spectral lines. The position of these lines for a large number of gases are stored in large databases known as spectroscopic databases. These databases are a considerable source of error in the retrieving methane concentrations in the atmosphere, due to the uncertainty of the position and the magnitude of these spectral lines. Differences in the various available spectroscopic databases could lead to significant differences between satellite estimates of methane. Understanding these differences is an important step towards reducing these uncertainties in future satellite measurements.

b) L44-50 should be at the end of the introduction, detailing the content of the paper. Up until that point, methane isotopologues have not been mentioned at all, maybe add a line about this aspect of the work as well beforehand.

We have now moved this section to the end, and we have introduced the concept of isotopologues further up in the introduction text.

c) I am not certain of the necessity of Table 1. One could either have a table with more information (a graph of the respective instruments and their spectral coverage), or a sentence in the text? Since you discuss TCCON windows shortly thereafter, maybe it would be a good opportunity to describe the TCCON windows used for the CH4 retrieval there (instead of referring to Table 4)?

We have now removed Table 1, and describe the spectral ranges, we think the original spectral fits, and as you identify, the description of TCCON windows in Table 4 (now Table 3) is sufficient to describe the overlap of TROPOMI/UVNS and TCCON. We also introduce the TCCON methane windows in this section.

d) L62-66: Maybe I misunderstood the point being made, but it seems simplistic to suggest that the retrieval window differences between TCCON and TROPOMI may be the major source of biases in the validation of the methane product, especially when many other factors will affect the quality of the retrieval on both GB and satellite platforms. This is surely an interesting point to study, but this seems to inflate the importance of the retrieval window on the validation.

Thank you for this point, however, we do not feel that this section identifies windows differences as the main source of error between TCCON and TROPOMI. We are attempting to highlight that the 4190-4340 cm-1 is a relatively unexplored region for methane, especially for TCCON, and that we can expect differences because of this. Indeed, one of the GGG developers identified that this spectral region has not been explored in TCCON historically because the spectroscopic parameters were not reliable.

e) L71: “We can infer some of the potential spectroscopic related biases in satellite retrievals…”: yes, but since the resolution of TCCON and satellite instruments are quite different, TCCON should be much more affected than satellite retrievals, maybe something to mention.

Agreed, we have added the following sentence.

Note that the spectral resolution of TCCON is typically significantly higher than that of TROPOMI and other satellite instruments, which are unlikely to affected to the same degree as TCCON.

e) L75: I think that Eq. 2 should be moved here. [?]

Agreed, This equation has been moved here.

3/ L119: “GFIT assumes a fixed profile shape for each trace gas, …”: you should mention how the a priori is determined

We have now identified the source of the a priori data in the sentence.

4/ L198-201: Is there already an estimate of the errors due to these assumptions? It would be interesting to have some figures on this in the paper.

This is an important point, the original paper describing TCCON (Wunch et al., 2011) does give these estimates in Figure 7. However they are only available for CO2, and are not shown for methane.

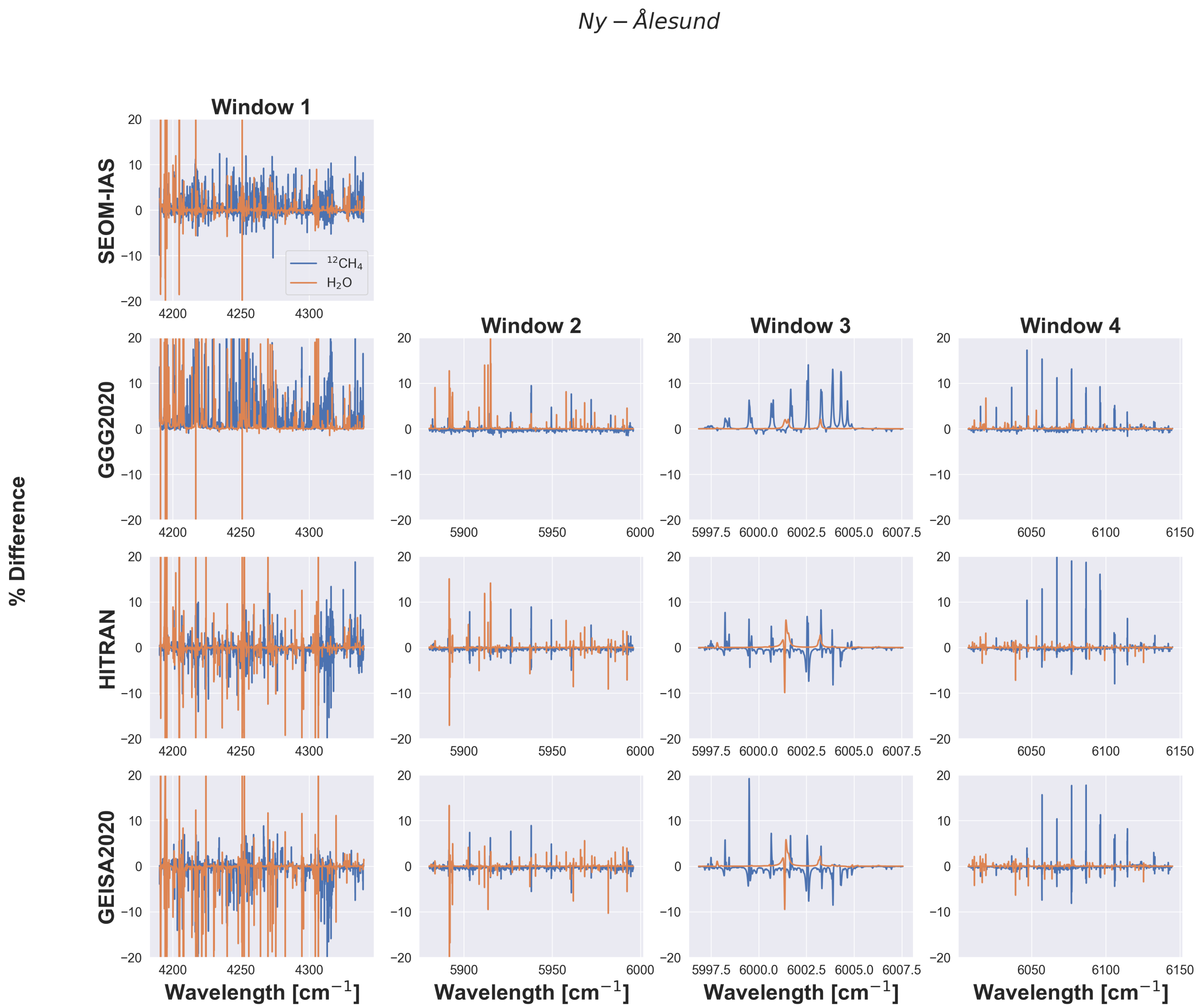
5/ Section 3 Before analysing the retrieval data, I would *suggest* to first perform a comparison of synthetic spectra using fixed state vectors with different spectroscopic databases to look at the expected impact in terms of transmissions in various windows. It might also be interesting to look at the contribution of H2O and CH4 to the total transmission, and hence have a better feeling for which windows should be more sensitive to the water content (would help with the analysis of section3.3?)

Thank you for the suggestions in this comment, firstly regarding the idea to investigate synthetic spectra, although this would be interesting, we feel the resources required to adequately describe and analyse the results, for a series of different locations (which is important in terms of the impact of a priori) would be significant.

We have added some figures to analyse how the calculated transmission of 12CH4 and H2O impact the total transmission. For example the figure below shows the % difference in calculated

12CH4 transmission w.r.t. GGG2014. We can see across all bands there are a number of differences, notably in window 1. We have also included an similar assessment for water vapour in the update to the paper. Comparing the figure below to Figure 1 in the original paper, we can see some of the large deviations matching in both figures.

We have updated section 3.1 to include an assessment on the specific transmission differences.



6/ Section 3.1 A slightly more exhaustive analysis of these results would be interesting. Could you associate some of the more prominent residuals with species? For instance, in window 3, HITRAN and GEISA have large discrepancies around 6001 cm-1 compared with GGG, why is that? Are there clear improvements due to the use of non-Voigt parameters?

Please see our response to the point above, where we tackle this issue. Specifically for the 6001 cm-1, we see (according to the figure above) that water vapour differences seem to be the main cause for the discrepancies. The figure above and the other analyses we have done, do not give obvious weight to the improvements of the non-Voigt parameters used in the SEOM-IAS database, apart from the constant lower magnitude RMSE values observed for SEOM-IAS database in Figure 2. Likely because the 4200-4400 cm-1 spectral region is highly complex, and cannot be directly compared with the 5900-6150 cm-1 spectral region. We can now attribute some of the more characteristic differences, such as the one you mention at 6001 cm-1 for HITRAN and GEISA to water vapour rather than methane.

7/ Section 3.3 In this part of the analysis, you sometimes seem to confuse correlation and sensitivity, which are quite different, or only consider one of these aspects. For instance, on L346 you state that "results from Ascension Island, which operates at lower SZAs than any of the other TCCON sites considered in this study also indicates large correlations, suggesting further complexity". The truth is that the sensitivity (slope) of the linear fit to SZA is close to 0, so despite a large correlation, very little effect is observed.

Thank you for this point, Reviewer #4 pointed out that some of the relationships may not be linear, so we have rewritten this section of the paper, and taken your comments into account. Clearly contrasting the difference between correlation and sensitivity.

On L348-349: "For window 1, there is weak sensitivity to water vapour variations at the Ny-Ålesund and the Darwin sites," where the sensitivity is substantial for Ny Alesund in Window 1, but correlation is weak.

This section has now been re-written.

Section 3.3 could also gain from an indication of statistical significance in the results. It is mentioned on L385 that large p-values were found in some cases, I would suggest identifying clearly results that are not significant, and state at what level you consider these not statistically significant.

Only the GGG2020 values for the Tsukuba site were found to be statistically insignificant according to the p-test. We have therefore removed the GGG2020 Tsukuba results from the updated figures (10, 11 and 12), and have identified in the caption and with a text box in the corresponding section that the p-test was failed for these results. While highlighting that the p-test was passed for all other results.

Typos and technical corrections
Comments with a "[?]" at the end are for your own consideration

L1: "methane retrieving satellites" -> "methane retrieving satellites instruments" [?]

Changed.

L3: "is as important as when TCCON…" -> "is as important now as when TCCON…"

Changed.

L6: the 'S' in "Spectroscopic databases" should not be capitalized

Changed.

L7: Leave out spectroscopic databases in "TCCON GGG2014 and GGG2020 spectroscopic databases;", as it is redundant [?]

Changed to "The spectroscopic databases include those native to TCCON GGG2014 and GGG2020"

L14: consider changing "~x3" by "up to three times" [?]

We changed this to "2.5", more in line with the results.

L16: "vapour. Such" -> the period should be a comma, and the Such should be lowercase. But you should consider changing this sentence as it is very long and uses 5 times the word "different"

The abstract has now been re-written.

L24: "bottom up" -> "bottom-up"

Corrected.

L25: "top down" -> "top-down"

Corrected.

L28: "S5P" -> "Sentinel 5-Precursor (S5P)" , "TROPOMI" -> "TROPOspheric Ozone Monitoring Instrument (TROPOMI)" . These acronyms have been defined in the abstract, but consider adding a proper acronym description in the full paper as well. [?]

Added.

L40: "will therefore be relying" -> "therefore rely"

Corrected.

L43: "from the 6000 cm-1 spectral range": a range has two values. Either define the window precisely, or reformulate

Changed to "spectral region".

L44: "make" -> "perform"

Corrected.

L48: "…quantify the variations in retrieval abundances when using five separate spectroscopic databases,…" : this seems like the main aspect of this paper, it should probably be mentioned first.

Thank you, we have now identified how errors can creep in due to different spectroscopic databases in the first paragraph.

L50: "are studies" -> "are studieD"

Thank you, corrected.

L50: consider a reformulation of this sentence

This sentence is now reformulated as:

Building on this assessment, the sensitivity of the retrievals to variations in water vapour concentration and path length are studied. This allows for the assessment of how differing windows and spectroscopic databases are sensitive to variations in local conditions.

L52: (an weighted …" -> "(a weighted…)

Corrected, thank you.

L81: Either state "98.9% and 1.1%", or "99% and 1%", but it should add up to 100%. Considering the "roughly" in the sentence, I would go with 99% and 1%.

Agreed, changed to 99% and 1%.

L86: "structured as follows," -> change the comma to a colon

Corrected to colon, thank you.

L90: "…conclusions are shown…" -> "conclusions are drawn" [?]

Agreed, changed.

L100: Add a comma after "In this study"

Added.

L100-101: "which is summarised briefly here." -> "briefly summarised here."

Changed, thank you.

L101: "A forward" -> lowercase "a"

Changed, thank you.

L104: “fit” -> “fitted” [?]

Changed, thank you.

L105: after “In the case of GFIT”, add a comma

Inserted, thank you.

L106-115: I am not convinced that these are essential to the papers and may be omitted. Maybe reconsider listing these, and put a reference instead [?]

We disagree with the reviewer here, these state vector elements are not typically listed, and are different from what is typically used in a satellite retrieval which we wish to emphasis. In addition, we wish to further emphasis the use of the ‘continuum curvature’ fit element which is not standard, even for TCCON.

L116: confusing wording: “Note that not all of the above are not routinely…” I imagine that you should omit the second “not”.

Good catch, thank you, the ‘not’ has been removed.

L116: Start a new sentence with “For example” [?]

We have replaced ‘for example’ with ‘especially’ which we think fits better.

L124: “the 7885 cm-1 spectral range multiplied” -> a spectral range needs 2 numbers. Maybe use a formulation such as “in the vicinity of 7885 cm-1…” [?]

Corrected to spectral region.

L127: Not sure if it’s necessary to state the resolution again, which was already mentioned on L57

Agreed, reference has been removed.

L131-132: not sure if this sentence needs commas after “window” and “study” or if they are optional [?]

Added a comma after “window”, thank you.

Table 4 has a beginning o f a sixth row

We are not sure why this happened, however we have changed the format of the table, and this additional row no longer appears.

L135: is there a reference for the “standardised process”?

We have included a reference to the TCCON wiki page, however this requires an account and log-on. There is no other available documentation for this process.

L137: this sentence should not be its own paragraph L158: “Further to exploring the…” -> maybe “Beyond exploring …” [?]

Agreed, changed to “Beyond exploring”.

L167: “line mixing” -> “line-mixing”

Corrected.

L169: “They find a 1.1% difference in total methane…” -> is it a positive or negative bias? Maybe just add a “+” to emphasize [?]

It’s a positive bias, a “+” has been added.

L184: “standard methane window” -> either “standard methane windows” (there are 3 of them, right?) or, better yet “standard TCCON CH4 product”

Agreed, changed to “standard TCCON CH4 product”.

L185-189: These are technically correct sentences but are a bit long and wordy. Maybe consider adding equations ? [?]

Agreed, we have replaced these sentences as equations.

L194: This sentence should not start a new paragraph, but simply continue the previous one.

We have removed the paragraph break.

L196: there should not be a period before the “2)”. Maybe a semi-colon?

The period has be replaced with a semi-colon.

L197-198: “apriori” or “a priori” -> be consistent, I have seen both in the paper

We double checked this, and are now consistent.

L202: “cross sections” -> “cross-sections”

Corrected.

L204-207: consider re-writing this sentence, it's a bit convoluted. Maybe with more than one sentence?

We have re-written this sentence as follows:

These dependencies are quantified by non-linear regression analysis, consisting of fitting the variations of water vapour, SZA and measured temperature against the normalised difference between each methane isotopologue DMF case and the DMFs from the standard TCCON methane retrieval window. Here the normalisation factor is the uncertainty from the standard TCCON methane retrieval.

L208-209: you probably should not hyphenate an acronym L

We don't see this acronym, TCCON is hyphenated, but likely because the text is wrapped around.

213: leave out "metric"

Removed.

L214: you repeat "varies in the total"

Thank you, removed.

L216: "<< This sentence is unclear (especially the second part: "through the fact…"). Also, it should be "long-term", not "long term".

This sentence has now been re-written as follows:

However, TCCON currently represents the best chance of remotely measuring d13C, since precision errors are low and SNR is high Wunch et al ., (2011). In addition, because TCCON sites are situated in fixed positions long-term averaging is possible, which further reduces precision based errors.

Figure 1: Would it be possible to plot or mention the instrumental uncertainty along the residuals?

Individual error contributions are not provided by GGG2014, and it is not possible to identify them until trace gas uncertainty is provided.

L238: "shown in Fig 2," -> "presented in Fig 2," [?]

Changed, thank you.

L239: "Fig 2 is the fit" -> "Fig 2 is that the fit"

Changed, thank you.

L247: Maybe start a new paragraph with “There are differences in the…” as you focus more on the TCCON sites now, not on the windows?

Agreed, changed.

L256: “Likely reasons” -> for what? If you refer to the last paragraph, you should stay in that paragraph, and state more clearly what you mean.

We placed this section listening the reasons into the previous paragraph, and rephrased to, Likely reasons for this difference are.

Figure 2: Maybe indicate in one plot (window4?) the number of spectra used for the statistics? I know this information is available in Table 2, but this is always nice to have this information on hand.

We don’t actually use all available data to generate the statistics in this plot (largely because this caused problems with my machine). We therefore generated these statistics from a subset of 500 measurements from each site, this is now indicated in the figure caption.

Figure 2 caption: maybe indicate the colour of each statistics in parenthesis: “Each subplot shows the RMSE (black) and chi2 values (blue) …”

Agreed, added.

L263+264: “fit” -> “fitted” [?]

Changed.

L263: “there no specific” -> “there are no specific”

Corrected.

L267: “GESIA” -> “GEISA”

Corrected.

L270: “the the” -> “that the”

Corrected.

L270: “retrievals each database” -> “retrievals for each database”

Corrected.

L274-275: This is not a complete sentence, there seems to be missing something (maybe Firstly we present the results?)

Good catch thank you, changed to “Firstly we analyse the results”

L289: “from in” -> delete “in”

Deleted.

L290: “the standard in” -> “the standard retrieval in”

Changed.

L304: “sides” -> “sites”

Corrected.

L327-330: This sentence encapsulates nicely the differing atmospheric conditions of the various TCCON stations. Maybe it should be moved to the introduction to better explain this part of the work [?]

Agreed, we have moved this sentence to section 2.1.

Figure 8: This figure is not very easy to read, but I am not quite sure how to improve it. It does give a good overview of the results. Maybe instead it could be a pcolor-type graph for the slope of the linear fit and the intercept for each station and parameter sensitivity (each row of the graph could be a different spectroscopic database, each column a different window?). Just a suggestion

We agree that the figure is challenging, however we think the following figure (now figure 10) captures the type of information the referee suggests here. We feel the point of figure 8 (now figure 9) gives a qualitative impression of the sensitivity of each database/window to variations in local conditions, and therefore remains useful. Especially now non-linear regression is used for fitting models, we see clear differences between non-linear sensitivities and linear sensitivities.

Figure 8 legend: “GEISA”, not “GESIA”

Corrected.

L334: The description of Figure 8 is really not clear. Could you be more precise in what is being shown in the figure instead of “the qualitative distributions”?

We have added the following text to explore this figure in more detail.

Figure 9 qualitatively describes the nature of the sensitivity of each TCCON site/database/window to variations in local conditions. For Ny-Ålesund (row 1), there is a mixture of non-linear and linear sensitivities to variations in water vapour and SZA. Windows 2, 3 and 4 for GEISA2019 indicate particularly significant non-linear sensitivities to SZA variations. Sensitivities to temperature variation are generally linear, although some indications of slight non-linear behaviour are apparent (GEISA2020). For Ny-Ålesund there are some cases where little sensitivity is observed, e.g. HITRAN, suggesting a wide range of responses in the databases/windows. In contrast to Ny-Ålesund, Darwin (row 2) shows limited sensitivity to local condition variations, with low magnitude linear gradients observed for most cases. There are some exceptions, notably HITRAN window 3 and GEISA2020 windows 3 and 4 in relation to SZA variations, were significant non-linear behaviour is observed. Tsukuba (row 3) again shows significantly different behaviour, with almost all databases/windows showing significant linear or non-linear sensitivity. Window 1 for SEOM-IAS, GGG2020, HITRAN and GEISA indicate significant negative linear relationships, with all other cases show a range of sensitivity. For variations in SZA, as with Ny-Ålesund and Darwin, HITRAN window 3 and GEISA2020 windows 3 and 4 suggest strong non-linear sensitivity to variations in SZA. Most of the other windows/databases indicate some linear/non-linear sensitivity, but not to the same degree as HITRAN window 3 and GEISA2020 windows 3 and 4. Temperature variations for Tsukuba indicate significant non-linear sensitivity for window 1 in most cases (except GGG2014), and in general show different results from those shown in Ny-Ålesund and Darwin. Finally for Ascension Island, we note almost no sensitivity to any local condition variation, except for HITRAN window 3 and GEISA2020 windows 3 and 4 with SZA variations, which have shown sensitivity in all cases.

L342: “indicate the retrieval” -> “indicate that the retrieval”

We have changed this section of the text, and this does not exist in the new text.

L345: “creep” -> maybe find a better word [?]

Changed to propagate.

L349: “opposite databases” -> what are opposite databases?!

This has been removed from the text.

L362: “vapour. Thus” -> should be in the same sentence “vapour, thus”

These words have been removed from the text in the updated paper.

L396: “in the calculation” -> its more in the analysis of the results than the actual calculation of the value…

Agreed, calculation has been changed to analysis.

Table 5: The table should be slightly improved in such a way that there is less text. Maybe use subrows for each station, one for each database (basically like it’s done now, but just mention the database name once, not in each cell…). Also, would it be possible to either add the uncertainty

on the d13C value, or mention an order of magnitude in the text? This would add important information to the discussion of L398-401 about the uncertainty.

The table has now been split into two, and modified to include less text.
W.r.t the uncertainty, we discuss the retrieved uncertainty of 13CH4 in more detail, and come to the conclusion the precision errors are not the limiting factor in calculating d13C. The sentences indicated have been updated, as shown below.

Tables 4 and 5 indicates a wide range of results, suggesting either significant differences in spectroscopic parameters or large retrieval uncertainty. GGG identifies the mean uncertainty of 13CH4 retrievals to between 0.5 - 2 ppb (~2.5-10%) depending on the database and TCCON site. However, given that these uncertainties can be averaged over a long period of time, they should reduce significantly (by $\sim$ x200 in the case of Darwin), meaning that the precision of 13CH4 retrievals should be very high (e.g. <0.006 ppb). Therefore precision errors cannot explain the differences in d13C values shown in Tables 4 and 5, meaning differences in the spectroscopic databases are the key sources of errors in 13CH4 retrievals. This therefore suggests that knowledge of 13CH4 retrievals spectroscopic parameters must be improved before serious attempts at remote sensing of 13CH4 can be made.

L398: “that we calculate the mean uncertainty” -> “that the mean uncertainty…is between…”

Changed.

L404: “show results” -> “are”

Corrected.

L407: “combination” -> “combinations” [?]

Corrected.

L420: “Further to this” -> “Furthermore” [?]

Corrected.

L447: “question that is of some interest to the community, “can we calculate realistic and constant 13C values from TCCON”.” -> this formulation is a bit strange to me, maybe use colon after community (instead of a comma), or change the formulation to something like “question that is of some interest to the community, namely whether it is possible to calculate realistic and constant 13C values from TCCON.”

We have reformulated this sentence as follows.

this study touches on a question that is of some interest to the community, namely whether it is possible to calculate realistic and constant d13C values from TCCON.

L453: “calculate” -> “calculated”

Corrected.

L459: “this is an assumption based” -> “it is based”

Corrected.

References:

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., & Wennberg, P. O. (2011). The Total Carbon Column Observing Network. *Phil. Trans. R. Soc. A*, *369*, 2087–2112. https://doi.org/10.1098/rsta.2010.0240

On the consistency of methane retrievals using TCCON and multiple spectroscopic databases.

Edward Malina[1], Ben Veihelmann[1], Matthias Buschmann[2], Nicholas M. Deutscher[3], Dietrich G. Feist[4,5,6], and Isamu Morino[7]

[1]Earth and Mission Science Division, ESA/ESTEC, Keplerlaan 1, Noordwijk, the Netherlands.
[2]Institute of Environmental Physics, University of Bremen, Bremen, Germany.
[3]School of Earth, Atmospheric, and Life Sciences, Faculty of Science, Medicine and Health, University of Wollongong NSW 2522 Australia.
[4]Lehrstuhl für Physik der Atmosphäre, Ludwig-Maximilians-Universität München, Munich, Germany.
[5]Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany.
[6]Max Planck Institute for Biogeochemistry, Jena, Germany.
[7]Satellite Remote Sensing Section and Satellite Observation Center, Center for Global Environmental Research, National Institute for Environmental Studies, Onogawa 15-2, Tsukuba, Japan.

Correspondence: Edward Malina (edward.malina.13@alumni.ucl.ac.uk)

Abstract. The next and current generation of methane retrieving ~~satellites~~ satellite instruments are reliant on the Total Carbon Column Observing Network (TCCON) ~~and other similar systems for validation, and understanding the biases between satellite and TCCON~~ for validation. Understanding the biases inherent in TCCON and satellite methane retrievals is as important now as when TCCON started in 2004. In this study we highlight possible biases between different methane products by assessing the retrievals of the main methane isotopologue $^{12}CH_4$.

Using the TCCON GGG2014 retrieval environment, retrievals are performed using five separate spectroscopic databases from four separate TCCON sites (namely, Ascension Island, Ny-Ålesund, Darwin and Tsukuba) ~~. The Spectroscopic databases include the~~ over the course of a year. The spectroscopic databases include those native to TCCON GGG2014 and GGG2020~~spectroscopic databases~~; the HIgh-resolution TRANsmission molecular absorption database 2016 (HITRAN2016); the Gestion et Etude des Informations Spectroscopiques Atmosphériques 2020 (GEISA2020) database; and the ESA Scientific Exploitation of Operational Missions - Improved Atmospheric Spectroscopy (SEOM-IAS) database. We assess the biases in retrieving methane using the standard TCCON windows and the methane window used by the Sentinel 5-Precursor (S5P) TROPOspheric Ozone Monitoring Instrument (TROPOMI) for each of the different spectroscopic databases.

By assessing the retrieved $^{12}CH_4$ values from individual windows against the standard TCCON retrievals, we find bias values of between 0.05 to 2.5 times the retrieval noise limit. These values vary depending on the window and TCCON site, with Ascension Island showing the lowest biases (typically < 0.5) and Ny-Ålesund or Tsukuba showing the largest. For the spectroscopic databases, GEISA2020 shows the largest biases, often greater than 1.5 across the TCCON sites, and considered windows. The TROPOMI spectral window (4190-4340 cm^{-1}) shows the largest biases of all the spectral windows, typically >1, for all spectroscopic databases, suggesting further improvements in spectroscopic parameters are necessary. We further

assess the sensitivity of these biases to locally changing atmospheric conditions such as solar zenith angle (SZA), water vapour and temperature.

~~We find significant biases when compared to standard TCCON retrievals, in some cases up to ~x3 the retrieval noise limit depending on the window and database. We also find strong evidence that different windows in different spectroscopic databases exhibit different levels of sensitivityto changing local conditions such as light path length and water vapour. Such that inter-comparisons between different instruments using different retrieval windows should take these sensitivities into account. Based on cross-comparison studies with the standard TCCON methane windows, retrievals using the S5P/TROPOMI spectral range show results as reliable as the operational TCCON products.~~ We ~~therefore recommend that this TROPOMI methane window should be considered in future TCCON methane retrievals~~find evidence of significant non-linear relationships between the variation of local conditions and the retrieval biases based on regression analysis. In general, each site/database/window combination indicating differing degrees of sensitivity, with GEISA2020 often showing the most sensitivity for all TCCON sites. Ny-Ålesund and Tsukuba shows the most sensitivity to local conditions variations, while Ascension Island indicates limited sensitivity.

Finally, we investigate the biases associated with retrieving $^{13}CH_4$ from each TCCON site and spectroscopic database, through the calculation of the $\delta^{13}C$ value. With the aim of assessing the consistency of $^{13}CH_4$ across the databases. We find high levels of inconsistency, in some cases >1000% between databases, suggesting more work is required to refine the spectroscopic parameters of $^{13}CH_4$.

1 Introduction

Methane is widely acknowledged to have a significant impact on the global climate (IPCC, 2014), but the processes via which it enters and is removed from the atmosphere are still poorly understood, with ~~bottom up~~ bottom-up (scaled up in-situ measurements) estimations of the global methane budget not agreeing with ~~top down~~ top-down estimations (models) (Kirschke et al., 2013; Saunois et al., 2019). This disconnect is one of many reasons that has led to the development of multiple satellite missions, with the aim of improving the knowledge of the global methane budget. The remote sensing of methane is fundamentally dependent on inferring atmospheric concentrations from the absorption of light at wavelengths unique to methane, otherwise known as spectral lines. Methane, like all gases, is composed of a number of isotopologues, for example $^{12}CH_4$ and $^{13}CH_4$ forming the main constituents of methane. The position and intensity of the spectral lines of these isotopologues are stored in large databases known as spectroscopic databases (Gordon et al., 2017). These databases are a considerable source of error in the retrieval of atmospheric methane abundances, due to the uncertainty of the position and the magnitude of these spectral lines. The uncertainty is less with more abundant isotopologues (for example $^{12}CH_4$), however rarer isotopologues (e.g.$^{13}CH_4$) can have far more uncertainty. Differences in the various available spectroscopic databases could lead to significant differences between satellite estimates of methane (Galli et al., 2012; Scheepmaker et al., 2016). Understanding the spectroscopic differences of methane isotopologues is an important step towards reducing these uncertainties in future satellite measurements, and further refine the databases.

The launch of the Sentinel 5-Precursor (S5P) satellite, with the ~~TROPOMI~~ TROPOspheric Monitoring Instrument (TROPOMI) instrument (Veefkind et al., 2012), and the future Sentinel 5 (S5) mission with its Ultra-Violet Near infrared Shortwave infrared (UVNS) instrument (Ingmann et al., 2012), represent a significant advancement in space-based Greenhouse Gas (GHG) remote sensing, building on a decade of progress from the Greenhouse Gases Observing Satellite (GOSAT) (Yoshida et al., 2013). Unlike GOSAT, TROPOMI and UVNS exploit the 4190 – 4340 cm^{-1} spectral range, which has not been explored in detail from previous space-based instruments for methane retrievals. The Scanning Imaging Absorption spectrometer for Atmospheric CartograpHY (SCIAMACHY) (Bovensmann et al., 1999) onboard the ENVIronmental SATellite (ENVISAT) was sensitive to this spectral range, but was plagued with detector issues (ice build-up). The Measurements Of Pollution In The Troposphere (MOPITT) instrument (Drummond and Mand, 1996) is also sensitive to this spectral range, but is also affected by technical issues and has never successfully retrieved methane in this spectral window. The follow-on to GOSAT (GOSAT-2) also uses this spectral range; processing for GOSAT-2 is currently on-going. In addition, the wide spectral sensitivity of the limb viewing Canadian Atmospheric Chemistry Experiment (ACE)- Fourier Transform Spectrometer (FTS) (Bernath et al., 2005) includes this spectral window, but again the methane products of ACE-FTS do not include retrievals in this window. S5P/TROPOMI and S5/UVNS ~~will therefore be relying~~ therefore rely on spectroscopic parameters for which only limited experience is available in their application to space-based methane retrieval instruments (Checa-Garcia et al., 2015; Galli et al., 2012). TCCON, although sensitive to this spectral range, has primarily provided its methane abundances retrieved from the 6000 cm^{-1} spectral ~~range~~region, allowing for direct comparisons with SCIAMACHY and GOSAT.

TCCON is a global network of 27 ground based Fourier Transform Spectrometers (FTS) (Wunch et al., 2010), with the primary aim of providing reference total column (~~an~~ a weighted average value for a nadir viewing profile) abundances of numerous atmospheric species calibrated against aircraft profiles (Wunch et al., 2010, 2011), including methane, for validation and cross-calibration purposes. TCCON operates in a wide spectral range (4000 – 11000 cm^{-1}) and records direct solar spectra. TCCON is currently one of the key sources of reference data for the validation of satellite-based GHG retrievals, e.g. the Orbiting Carbon Observatory (OCO)-2, GOSAT and TROPOMI (Yoshida et al., 2011; Crisp et al., 2012; Lorente et al., 2021). TCCON instruments have both high spectral resolution (0.02 cm^{-1}), and high Signal to Noise Ratios (SNR) due to direct solar viewing geometry, and insensitivity to atmospheric scattering, thus making TCCON measurements higher quality than satellite measurements and excellent comparison datasets for satellite retrievals.

~~TCCON and TROPOMI /~~TROPOMI and UVNS both have overlapping spectral windows ~~in the~~ with the wide spectral range of TCCON within the Shortwave Infrared (SWIR) methane absorption regions, ~~highlighted in Table 1.~~

~~Methane SWIR windows commonality between S5P/S5 and TCCON Methane window S5P/TROPOMI S5/UVNS TCCON~~ 5970-6289 cm^{-1} ~~N Y Y~~ for UVNS and 4190-4340 cm^{-1} ~~Y Y Y~~ for UVNS and TROPOMI.

When validating methane products from TROPOMI and UVNS, retrieval products using the 4190 – 4340 cm^{-1} window will be compared with TCCON methane products generated using the standard TCCON windows ~~(Table 3)~~ 1) 5880-5996 cm^{-1}, 2)

5996.45-6007.55 cm^{-1} and 3) 6007-6145 cm^{-1}. Therefore potential biases associated with the choice of fit windows should be quantified and understood. Indeed, if the 4190 – 4340 cm^{-1} window proves to be as accurate as the standard TCCON windows, then there is justification to integrate TCCON retrievals from this window into future TCCON retrieval products. In addition, numerous algorithms will be used to provide methane data products from TROPOMI/UVNS (Hu et al., 2016; Schneising et al., 2019), which may use differing spectroscopic databases and are therefore subject to differing biases. Building on examples of similar past studies (Checa-Garcia et al., 2015; Galli et al., 2012), the high SNR and high spectral resolution makes TCCON data an excellent resource to assess any potential variations due to differences in the spectroscopic databases. By investigating the biases present in TCCON observations made at several sites over several seasons. We can infer some of the potential spectroscopic related biases in satellite retrievals, and their dependencies on local conditions such as water vapour that are relevant to ongoing TROPOMI validation, and future S5/UVNS validation. Note that the spectral resolution of TCCON is typically significantly higher than that of TROPOMI and other satellite instruments, which are unlikely to affected to the same degree as TCCON.

In addition to assessing the window and spectroscopic source biases for the main methane ~~isotopologues~~isotopologue $^{12}CH_4$, the opportunity is taken to retrieve the second most abundant isotopologue $^{13}CH_4$, and from this calculate the $\delta^{13}C$ ~~metric~~ value (see Eq 1). $\delta^{13}C$ requires the concentration of the two main methane isotopologues $^{12}CH_4$ and $^{13}CH_4$, which make up roughly 99% and 1% of global atmospheric methane respectively. Almost all measurements of this value are limited to in situ studies or airborne flask measurements, which although highly accurate, by their nature are spatially limited. Some effort has gone into satellite based retrievals of this metric (Buzan et al., 2016; Weidmann et al., 2017; Malina et al., 2018, 2019), but the results of these studies show this to be a challenging task. Therefore the calculation of the $\delta^{13}C$ value is a target of secondary importance in this study.

$$\delta^{13}C = \left(\frac{(^{13}CH_4/^{12}CH_4)_{sample}}{(^{13}CH_4/^{12}CH_4)_{VPDB}} - 1\right) \times 1000\text{‰}, \qquad (1)$$

where VPDB refers to Vienna Pee Dee Belemnite, an international reference standard for ^{13}C assessment. This is a ~~metric~~ quantity that has been used in numerous studies globally to differentiate methane source types (Fisher et al., 2017; Nisbet et al., 2016; Rigby et al., 2017; Rella et al., 2015), e.g. fossil fuel burning or wetlands. Tropospheric methane typically exhibits a $\delta^{13}C$ value of roughly -47‰ (Rigby et al., 2017), and total column measurements from TCCON should not deviate from this value to a significant degree. Therefore this tropospheric $\delta^{13}C$ value acts as a useful proxy, to determine the stability and variability associated with retrievals of methane isotopologues from different spectral windows, spectroscopic databases, location and time using the tropospheric $\delta^{13}C$ value as a baseline. In terms of $^{13}CH_4$, there are no published precision and accuracy requirements or statistics with TCCON. Calculating total column values of this metric would be highly beneficial for understanding the global methane budget, but is unlikely to be achievable with TCCON with an accuracy that would be sufficient for that purpose. However, calculation of $\delta^{13}C$ with TCCON will allow for an assessment of how far current technology is from making a useful total column assessment. ~~$\delta^{13}C$ metric requires the concentration of the two main methane isotopologues $^{12}CH_4$ and $^{13}CH_4$, which make up roughly 99% and 1.1% of global atmospheric methane~~

~~respectively. Almost all measurements of this metric are limited to in situ studies or airborne flask measurements, which although highly accurate, by their nature are spatially limited. Some effort has gone into satellite based retrievals of this metric (Buzan et al., 2016; Weidmann et al., 2017; Malina et al., 2018, 2019), but the results of these studies show this to be a challenging task. Therefore the calculation of~~

In this study, we use the TCCON GGG2014 (Toon, 2015) environment as the main tool for retrievals. Spectra are taken from four different TCCON sites in order to assess the impact of varying atmospheric conditions at different global locations. We assess the differences in abundances of the isotopologues and the quality of the fits when retrieved from standard TCCON spectral windows, and methane spectral windows in the TROPOMI/UVNS spectral range. We also quantify the variations in retrieval abundances when using five separate spectroscopic databases, and the application of non-Voigt line broadening shapes. Building on this assessment, the ~~δ^{13}C metric is a target of secondary importance in this study~~sensitivity of the retrievals to variations in water vapour concentration and path length are studied. This allows for the assessment of how differing windows and spectroscopic databases are sensitive to variations in local conditions.

This paper is structured as follows~~,~~ : section 2 outlines the methods used in this study, including details about the TCCON sites and spectra used, as well as the retrieval method. Information about the spectroscopic databases used in this study are also given. The results of this study are shown in section 3 outlining the biases between sites and databases, including an assessment of the sensitivity of the retrievals to local condition variability. Section 4 discusses the results shown in sections 3, and conclusions are ~~shown~~ drawn in section 5.

2 Methods, tools, datasets and requirements

2.1 TCCON sites used in study

We use TCCON spectra from four different sites identified in Table 1. Datasets over a single year were chosen in order to represent a wide range of seasonal conditions. The years chosen represent the years with maximum data coverage for each site respectively.

Table 1. TCCON sites used in this study.

TCCON Site	Lat/Lon	Date Range	Number of Spectra	Conditions
Ascension Island, Atlantic ocean	7.92°S, 14.3°E	Jan-Dec 2015	1518	Arid, little precipitation subject to some seasonal variation.
Darwin, Australia	12.5°S, 130.9°E	Jan-Dec 2020	39160	Tropical, significant water vapour background.
Ny-Ålesund, Spitsbergen	78.9°N, 11.9°E	April-Oct 2019	6315	Cold, dry, limited short-term variability.
Tsukuba, Japan	36.1°N, 140.1°E	Jan-Dec 2020	6162	Seasonal, cold dry winters, hot wet summers.

The TCCON sites used in this study were picked to have a wide range of conditions, with Ny-Ålesund capturing spectra in largely unvarying conditions with high SZA and low water vapour, while Ascension Island is similar in unvarying conditions although with higher background water vapour conditions and lower SZAs. This is contrasted by Darwin and Tsukuba which capture spectra under a wide range of SZAs and highly variable water vapour conditions. The mean background conditions for each site, as well as the variations over the dataset periods shown in Table 1 are indicated in Table 2. Significant variations in conditions and SZA are apparent between the TCCON sites, suggesting a wide range of capture conditions. We note the distributions of the conditions shown in Table 2 may not be normally distributed, but these statistics serve as a useful baseline to show the condition variations between the sites.

Table 2. TCCON sites water vapour, temperature and SZA average and variation.

TCCON Site	Water Vapour mean $\pm\sigma$	Temp mean $\pm\sigma$	SZA mean $\pm\sigma$
Ascension Island	4510 ppmv $\pm$ 890	27.5 °$\pm$ 1	38°$\pm$ 18
Darwin, Australia	5430 ppmv $\pm$ 1740	30.9°$\pm$ 3	45°$\pm$ 18
Ny-Ålesund, Spitsbergen	1440 ppmv $\pm$ 600	1.7°$\pm$ 6	69°$\pm$ 8
Tsukuba, Japan	3200 ppmv $\pm$ 2470	22.9°$\pm$ 9	50°$\pm$ 18

2.2 GFIT Retrieval Algorithm

In this study, we use the GGG2014 environment, which includes the GFIT retrieval algorithm (Wunch et al., 2010), ~~which is~~ summarised briefly here. GFIT employs a nonlinear least-squares fitting scheme: ~~A~~ a forward model (radiative transfer model which simulates radiation transfer through an atmosphere or a body of gas) is used to calculate synthetic irradiance spectra based on a set of parameters known as state vector elements (typically trace gas concentrations) and model parameters (e.g. temperature and pressure profiles). These synthetic irradiance spectra are then ~~fit~~ fitted to the measured irradiance spectra by adjusting the state vector elements to provide a final result, normally a trace gas abundance. In the case of GFIT, the state vector can include the following.

- first target gas scaling factor (desired output).
- interfering gas scaling factor.
- continuum level of the irradiance spectrum.
- continuum tilt
- continuum curvature
- frequency shift
- zero level offset
- solar scaling (differences in shifts of atmospheric and solar lines)
- fit channel fringes

Note that not all of the above are ~~not~~ routinely included in the state vector, ~~for example~~ especially the continuum curvature ~~especially~~ which is not commonly included in the state vector. This option is designed to remove instrument features, but may also attempt to remove other effects due to the spectroscopic database, as noted in the TCCON wiki (TCCON, 2020). GFIT assumes a fixed profile shape for each trace gas, and the sub-column amounts for each altitude/pressure level are not independently scaled. Unlike in most satellite retrieval algorithms, aerosol and albedo terms are not included in the state vector, because TCCON operates in direct solar viewing, where scattering is considered unimportant and surface terms are not necessary. The retrieved trace gas column is calculated by multiplying scaling factors from the retrieved state vector by the a priori vertical column abundances. The TCCON a priori profiles are obtained from the National Centers for Environmental Prediction/National Center for Atmospheric Research analyses for temperature, pressure and humidity. Combined with empirical models for CO_2, CO, CH_4 and N_2O developed from FTS balloon flights, data from the ACE-FTS instrument (Wunch et al., 2011). Dry air Mole Fractions (DMF) are calculated by dividing the scaled trace gas column with the total column O_2, retrieved from a wide window in the 7885 cm^{-1} spectral ~~range~~ region multiplied by the volume mixing ratio of O_2 0.2095. We use O_2 from GGG2014 only, to provide a point of consistency between the spectroscopic databases, and because SEOM-IAS does not include spectral lines in this region. DMF gas volumes identify retrieved abundances as mole fractions, as opposed to absolute concentrations, all retrieved $^{12}CH_4$ abundances are referred to as DMF values.

Because of the high spectral resolution of the TCCON instruments (0.02 cm^{-1}), most spectral lines are resolved, therefore radiative transfer calculations are performed on a line-by-line basis. GGG includes a spectroscopic database in its environment, which is similar to other more widely adopted databases~~(see below)~~. TCCON has a standard set of spectral windows for methane retrievals, all of which are in the 6000 cm^{-1} methane absorption window range. In this study we include the TROPOMI/UVNS SWIR spectral windows~~(4190-4340 cm^{-1})~~. This window, along with a description of all of the windows considered in this study are described in Table 3 below.

Table 3. Spectral windows used in study.

Window	Window spectral range (cm^{-1})	Target species	Background species	Window source
1	4190-4340	$^{12}CH_4$	CO_2, H_2O, HDO, CO, HF, N_2O, O_3	Sentinel 5 baseline
2	5880-5996	$^{12}CH_4$	CO_2, H_2O, N_2O	TCCON standard
3	5996.45-6007.55	$^{12}CH_4$	CO_2, H_2O, N_2O, HDO	TCCON standard
4	6007-6145	$^{12}CH_4$	CO_2, H_2O, N_2O, HDO	TCCON standard

Windows 2-4 are standard TCCON methane retrieval windows which in this study are used for $^{12}CH_4$, and window 1 is based on the TROPOMI spectral window (Galli et al., 2012; Hu et al., 2016), given that no standard windows exist in this spectral window for TCCON. In addition, TCCON methane products are the result of a standardised process where a weighted average of three retrieved values from windows 2, 3 and 4 described in Table 3 (Cal, 2022).

For $^{13}CH_4$ retrievals, windows 1 and 4 are used.

2.3 Spectroscopic Databases

We use parameters from five separate spectroscopic databases, which are as follows: 1) The database included with GGG2014 (Toon, 2015), which currently assumes a Voigt line shape for all lines. 2) The database included with the updated GGG2020 software, referred to in this study as GGG2020, which includes numerous updates to the GGG2014 spectroscopic parameters~~, and some application of~~. Some non-Voigt parameters are included in GGG2020, but are not exploited in this study, because GGG2014 was not modified to take advantage of them. 3) HITRAN2016, which is a well-established spectroscopic database that has been used in numerous satellite based studies previously (Galli et al., 2012). Methane has been updated in the current release HITRAN2016 (Gordon et al., 2017) from the previous release (HITRAN2012), with new lines and parameters included for both of the main isotopologues. HITRAN2016 includes the additional parameters required to model non-Voigt lines shapes, however the current version does not include these parameters for methane (at the time of writing). 4) The GEISA2020 database (Delahaye et al., 2021) is another spectroscopic database, similar in design and goals to the HITRAN databases. The GEISA database does not currently include non-Voigt line shape parameters. 5) SEOM-IAS (Birk et al., 2017), specifically developed for the TROPOMI spectral window and designed around non-Voigt atmospheric line shape profiles. This database only has data within the 4190-4340 cm^{-1} spectral range, and can therefore only contribute to window 1 of this study.

For clarification purposes, there are no official releases of the spectroscopic parameters used in the GGG TCCON retrievals. We refer to the databases used in this study as GGG2014 and GGG2020 in order to differentiate with them, based on the GGG retrieval environment releases, with GGG2020 due for release in the near future (Laughner et al., 2021).

Some work has been performed previously comparing spectroscopic databases e.g. (Jacquinet-Husson et al., 2016; Armante et al., 2016), generally indicating that the need to resolve differences between spectroscopic databases remains. Yet none have specifically targeted the TROPOMI SWIR spectral region, therefore this study is the first case with respect to the TROPOMI spectral window with TCCON.

~~Further to~~ Beyond exploring the impact of differing spectroscopic database parameters, we investigate the use of non-Voigt broadening parameters. Ngo et al. (2013) find the standard Voigt profiles used for spectral line broadening may be inadequate for trace gas retrievals (based on laboratory studies), which can lead to errors larger than instrument precision requirements. In order to calculate more accurate line shapes for remote sensing purposes, numerous models have been proposed. In this paper we use the quadratic Speed Dependent Hard Collision (qSDHC) model (Ngo et al., 2013; Tran et al., 2013). This model includes additional parameters based on speed dependence of collisional broadening and velocity changes of molecules due to collisions, on top of the standard parameters of pressure-induced air broadening, and pressure induced line shift. Note that only the SEOM-IAS database uses these additional parameters, the remaining spectroscopic databases do not include these parameters for methane at the time of this paper. We use the FORTRAN routines provided with Ngo et al. (2013) to implement the qSDHC model into the GFIT algorithm, modified to include first order Rosenkranz ~~line mixing~~ line-mixing effects. Mendonca et al. (2017) report that incorporating speed dependent and ~~line mixing~~ line-mixing has a significant effect on calculated methane columns when compared against assuming Voigt dependency. They find a +1.1% difference in total methane column abundances from 131,124 spectra. The implication is that it is important to account for the additional physical parameters included in non-Voigt models, when retrieving methane.

We note the introduction of $^{13}CH_4$ into spectroscopic databases in the TROPOMI spectral region is relatively recent, and in the case of HITRAN, was only introduced in the 2012 release (Brown et al., 2013), thus suggesting that $^{13}CH_4$ spectroscopic parameters may retain high levels of uncertainty.

2.4 Analysis structure and metrics

The following section describes the assessment metrics used in this study. Firstly we assess the quality of the fit of the measured and modelled spectra for each window indicated in Table 3 for each spectroscopic database at each TCCON site. The quality of the fit is expressed through Root Mean Square Error (RMSE) of the residual between the calculated transmission spectra, and the TCCON measurement transmission spectra, and the χ^2 test, quantitatively defined as:

$$\chi^2 = \sum_i [\mathbf{y_{measured}} - \mathbf{y_{calculated}}]^2. \tag{2}$$

.

Where $\mathbf{y_{measured}}$ refers to the measured TCCON spectrum, and $\mathbf{y_{calculated}}$ is the synthetic spectrum calculated by the forward model. Secondly we assess the variance of the calculated DMFs of $^{12}CH_4$ for each window, spectroscopic database and TCCON site w.r.t the standard methane ~~window~~ product used in TCCON retrievals currently, which is a weighted average of windows 2, 3 and 4. This variance is described through the RMSE ~~of the residual between~~ in Eq. 3:

$$NRMSE = \sqrt{\frac{\sum^{n}\left(\frac{X^{12}CH_{4window} - X^{12}CH_{4standard}}{\sigma_{standard}}\right)^2}{n}}, \tag{3}$$

where NRMSE refers to the normalised RMSE, $X^{12}CH_{4window}$ is the retrieved DMF ~~of $^{12}CH_4$ for~~ from a specific window~~and~~, $X^{12}CH_{4standard}$ is the retrieved DMF ~~of the standard TCCON retrieval window, normalised by dividing by~~ from the TCCON standard product, and $\sigma_{standard}$ is the retrieval error ~~of the standard TCCON retrieval window (NRMSE)~~from the standard methane product. The variance is also given by the absolute mean residual ~~between the retrieved DMF of $^{12}CH_4$ for a specific window and the retrieved DMF of the standard TCCON retrieval window, normalised by dividing by the retrieval error of the standard TCCON retrieval window (NAmean).~~ described in Eq. 4:

$$NAmean = \left|\frac{\sum^{n}\left(\frac{X^{12}CH_{4window} - X^{12}CH_{4standard}}{\sigma_{standard}}\right)}{n}\right|, \tag{4}$$

where NAmean is the normalised absolute mean residual, and all other terms are as identified previously. Following the assessment of the retrieval variance between windows and databases, we investigate if locally changing conditions impact biases between spectroscopic databases and windows. Variations in the retrieval conditions throughout the course of a day of measurements are included in TCCON error budgets, for example artefacts can appear in TCCON retrievals at extreme SZA values (Wunch et al., 2011).

We therefore investigate if the methane retrieval biases vary with respect to the following local parameters 1) SZA, where extreme angles can cause errors in the air-mass assumptions and affect characteristics of the Instrument Lineshape function (ILS) (Wunch et al., 2011)~~.~~; 2) Water vapour (retrieved by TCCON in two standard narrow windows) through the ~~course of a day at a range of~~ whole available data range at the respective TCCON sites. The GFIT retrieval algorithm is a scaling retrieval algorithm for all trace gas fitting, meaning that an incorrect ~~apriori~~ a priori trace gas profile shape will yield errors in the retrieval. The GFIT water vapour ~~apriori~~ a priori is based on a profile taken at midday for each specific retrieval, meaning that any significant variations from this daily profile will yield errors in the retrieval, that will vary depending on the impact of water vapour on a specific spectral window. 3) Temperature, which is not included in the retrieval state vector and dependencies on temperature will not be removed in the retrieval process. Temperature errors are introduced through the spectroscopic ~~cross sections~~ cross-sections (An et al., 2011), therefore poor knowledge of spectroscopic parameters will potentially lead to temperature based errors.

These dependencies are quantified by ~~identifying the possible existence of a linear correlation (using Pearson correlation coefficient and linear fit gradient) between~~ non-linear regression analysis, consisting of fitting the variations of water vapour, SZA and measured temperature against the ~~bias between the retrieved methane isotopologue DMFs for each window and spectroscopic database, against~~ normalised difference between each methane isotopologue DMF case and the DMFs from the standard TCCON methane retrieval window~~normalised by the noise~~. Here the normalisation factor is the uncertainty from the standard TCCON methane retrieval~~window~~.

The magnitude of the metrics defined above can be put into context by comparisons with the TCCON error budget. TCCON typically aims for precision of <0.3% on methane retrievals, and has a rough estimate of 1% systematic uncertainties (dominated by in-situ calibration which can affect sites differently (Wunch et al., 2015)). Therefore it is possible to judge the variation of $^{12}CH_4$ DMFs between windows and databases based on these biases and precision.

Finally, although the quality of the $^{13}CH_4$ fit metrics in this study are not covered in detail, we instead calculate ~~the~~ δ^{13}C ~~metric~~ in order to understand the plausibility and variation of retrieving $^{13}CH_4$ from TCCON. Fundamentally the final aim of retrieving $^{13}CH_4$ is to calculate δ^{13}C. How much δ^{13}C varies in the total ~~varies in the total~~ column is a complex issue (Weidmann et al., 2017; Malina et al., 2018, 2019), in-situ studies (Nisbet et al., 2016; Rigby et al., 2017; Fisher et al., 2017) all show that an uncertainty of «1‰ in δ^{13}C is required in order to determine natural annual variability at the surface. However, variability in δ^{13}C can be higher in the troposphere and stratosphere due to variability of the OH sink and the fractionation caused by OH (Röckmann et al., 2011; Buzan et al., 2016), with evidence that δ^{13}C can vary by up to 10‰ in different air parcels (Röckmann et al., 2011). Based on these factors, we assume a rough total column δ^{13}C variability of 1‰, which equates to a total uncertainty of <0.02 ppb on $^{13}CH_4$ retrievals, or roughly 0.1% of the total column. This is clearly an unrealistic target for individual retrievals, given the uncertainty requirements for $^{12}CH_4$ described above. ~~Nevertheless precision errors will be low due to the nature of TCCON, and through the fact that~~ However, TCCON currently represents the best chance of remotely measuring δ^{13}C, since precision errors are low and SNR is high (Wunch et al., 2011). In addition, because TCCON sites are situated in ~~a fixed position, allowing for long term averaging to reach a required precision target~~fixed positions long-term averaging is possible, which further reduces precision based errors. Therefore one of the minor aims of this study is to identify how far away TCCON uncertainty (including systematic errors) is from the desired uncertainty of <1‰ δ^{13}C.

$$\delta^{13}C = \left(\frac{(^{13}CH_4/^{12}CH_4)_{sample}}{(^{13}CH_4/^{12}CH_4)_{VPDB}} - 1 \right) \times 1000‰,$$

~~where VPDB refers to Vienna Pee Dee Belemnite, an international reference standard for ^{13}C assessment. Tropospheric methane typically exhibits a δ^{13}C value of roughly -47‰ (Rigby et al., 2017), and total column measurements from TCCON should not deviate from this value to a significant degree. Therefore this tropospheric δ^{13}C value acts as a useful proxy, to~~

~~determine the stability and variability associated with retrievals of methane isotopologues from different spectral windows, spectroscopic databases, location and time using the tropospheric δ^{13}C value as a baseline. In terms of $^{13}CH_4$, there are no published precision and accuracy requirements or statistics with TCCON.~~

3 Results

3.1 Quality of spectral fitting

An example of residual transmission spectra from the Ny-Ålesund site is shown in Fig 1, with the standard deviation of a selection of retrievals within the same time period indicated by the red lines. Examples of spectral fits from the other TCCON sites considered in the study are shown in the appendix. Qualitatively we note clear differences in the quality of the fits between windows and databases, for example there are clear deviations apparent, especially in window 1 for HITRAN and GEISA.

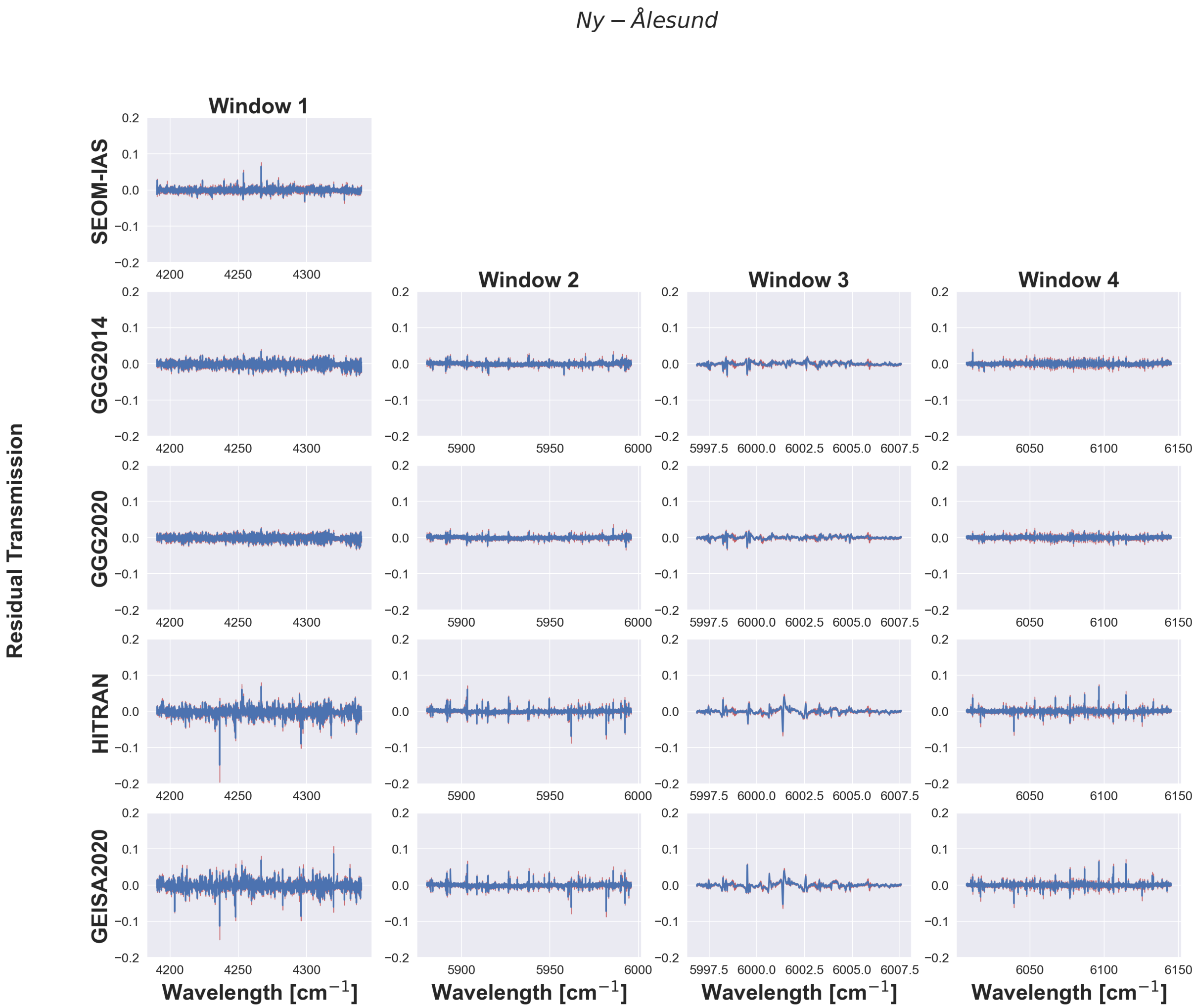


Figure 1. Example residual transmission spectra calculated from measured and fitted spectra from the Ny-Ålesund in 2019. The blue line indicates an example of the fit residual between the calculated transmission and the measured transmission. The red lines indicates the standard deviation of the residual, based on all spectra taken over the entire dataset. The columns of this figure identify the residuals of a specific window, and the rows a specific database, as identified in the axis labels.

The analysis statistics for the residual transmission spectra (as discussed in sect. 2.4) shown in Fig 1 are ~~shown~~ presented in Fig 2, as well as the associated statistics for the other TCCON sites considered in this study. What is clear from Fig 2 is that the fit statistics for each spectroscopic database, irrespective of TCCON site and window generally have the same pattern in terms of quality. For window 1 SEOM-IAS ~~usually has the best fit metrics (i.e. the lowest magnitudes)~~and GGG2020 are more or less equivalent in quality, followed by ~~GGG2020,~~ GGG2014, HITRAN and then GEISA. In windows 2-4 where SEOM-IAS has no data, GGG2020 typically shows the highest quality fits, suggesting the latest iteration of the GGG2020 spectroscopic parameters has superior performance to the older version.

Window 1 typically shows the poorest fit metrics of all windows, possibly because it is the largest window, but also because it is more complex region in terms of absorption (Brown et al., 2013) than the other windows. Window 4 for example is also wide, but typically shows higher quality fits than any of the other windows in this study. The implication being that the knowledge of spectroscopic parameters in window 1 is still lacking in comparison to the traditional TCCON windows.

There are differences in the metrics between TCCON sites, with Ascension Island showing poorer ~~fits~~ RMSE values than any of the other sites, similar to Darwin. This is to be expected however since these instruments are not identical, and capture spectra under differing conditions. We note that all instruments are run according to TCCON specifications but their respective configurations are not exactly the same. This is normal and necessary as different sites need local adjustments to account for different local conditions such as altitude, humidity or cloud conditions. Most of the effects caused by such individual configurations are removed by the differential CO_2 and CH_4 DMF retrievals but will affect individual spectra. For example, in the case of Tsukuba and Ascension, the configuration effects cannot be compared directly except for detector noise, which turned out to be comparable. However, the signal on the detector of the Ascension Island instrument is at least 50% lower than that of the Tsukuba instrument.

~~Likely reasons~~ Likely reasons for this difference are: 1) The Ascension FTS runs on a higher spectral resolution (0.014 cm^{-1} vs. 0.02 cm^{-1}) and a faster scanner speed (10 kHz vs. 7.5 kHz). Both reduce integration time per spectral pixel. 2) The illumination of the InGaAs detector on Ascension is kept low on purpose to avoid saturation. This setting cannot be readjusted in between site visits and has to last for months. Other sites may use similar techniques, and may vary depending upon need. 3) The solar tracker has known issues with pointing at the centre of the sun at low SZAs but cannot be replaced easily. In addition, dust buildup on the solar tracker mirrors reduces the reflectively of the mirrors quickly. They are cleaned weekly but a signal loss in the order of 20% over a few days is not uncommon.

The results from Tsukuba are different from the other showcased results, this is likely because of the smaller amount of spectra available for plotting (owing to data transfer and storage limits). There was no limit in the actual retrievals identified in the following paper sections.

Example transmission spectra for Darwin, Tsukuba and Ascension Island are identified in Appendix A.

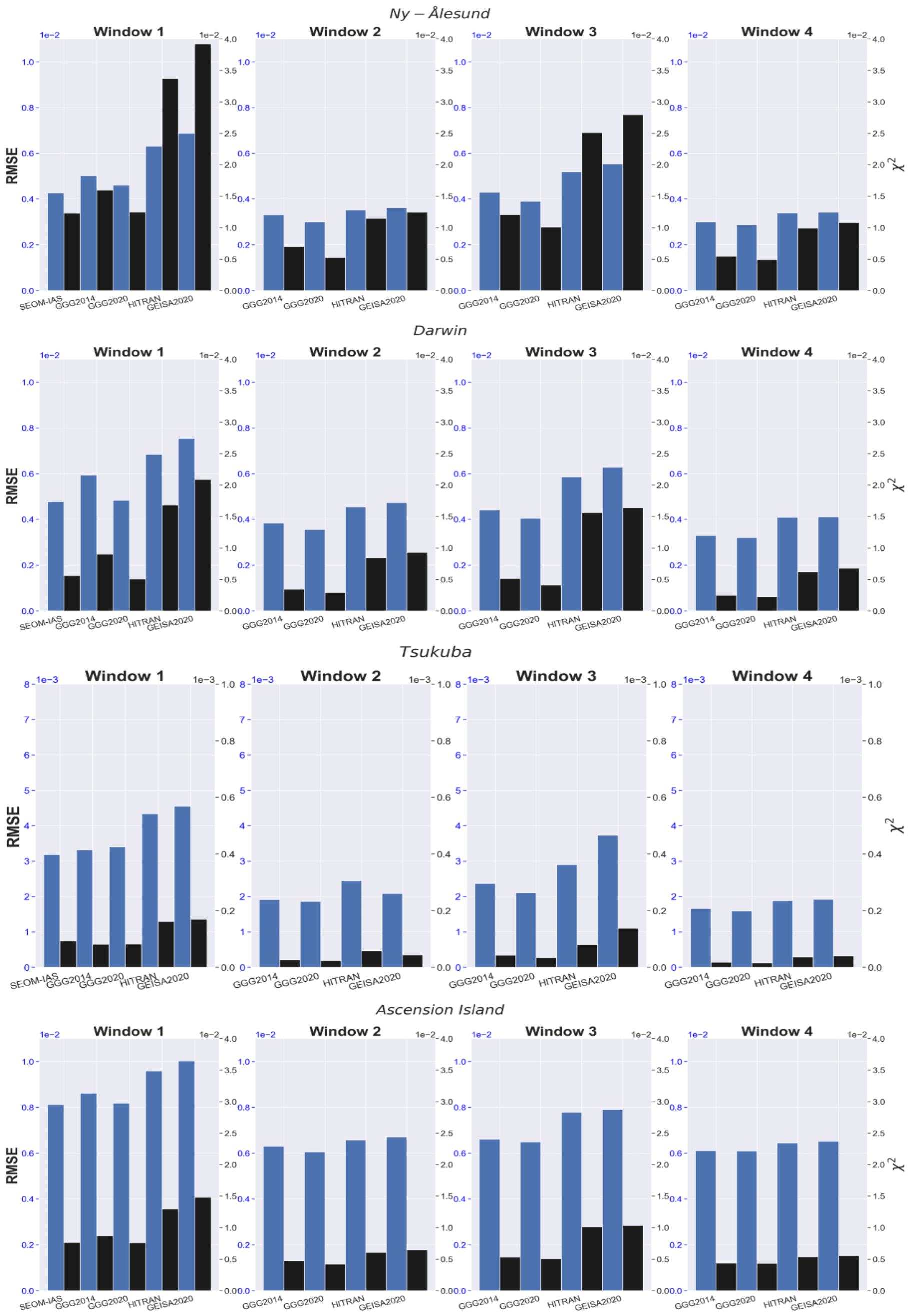


Figure 2. Bar chart indicating the fit statistics for a selection of retrievals from each of the TCCON site. Each row of the figure refers to results from each of the TCCON sites, indicated by the row title. Each column shows the results from each window, indicated by the title of each column. Each subplot shows the RMSE and χ^2 values for each spectroscopic database indicated in the x-axis, with the blue bars referring to the RMSE values, with magnitudes shown on the left-hand y-axis. The black bars refer to the χ^2 values, with the magnitudes indicated on the right-hand y-axis. Note the scale on the Tsukuba row is slightly different, to account for the lower magnitude results. A cross section of 500 spectra for each TCCON site are used to generate the statistics in this figure.

Since all trace gases are ~~fit~~ fitted simultaneously in all of the windows, there are no specific metrics associated with $^{13}CH_4$. $^{13}CH_4$ in this study is ~~fit~~ fitted in windows 1 and 4.

Building on the residuals indicated in Fig. 1, we investigate the differences observed in transmission residuals. We calculate the percentage difference for each spectroscopic database, w.r.t the tranmissions calculated by the GGG2014 database, for $^{12}CH_4$ and H_2O species.

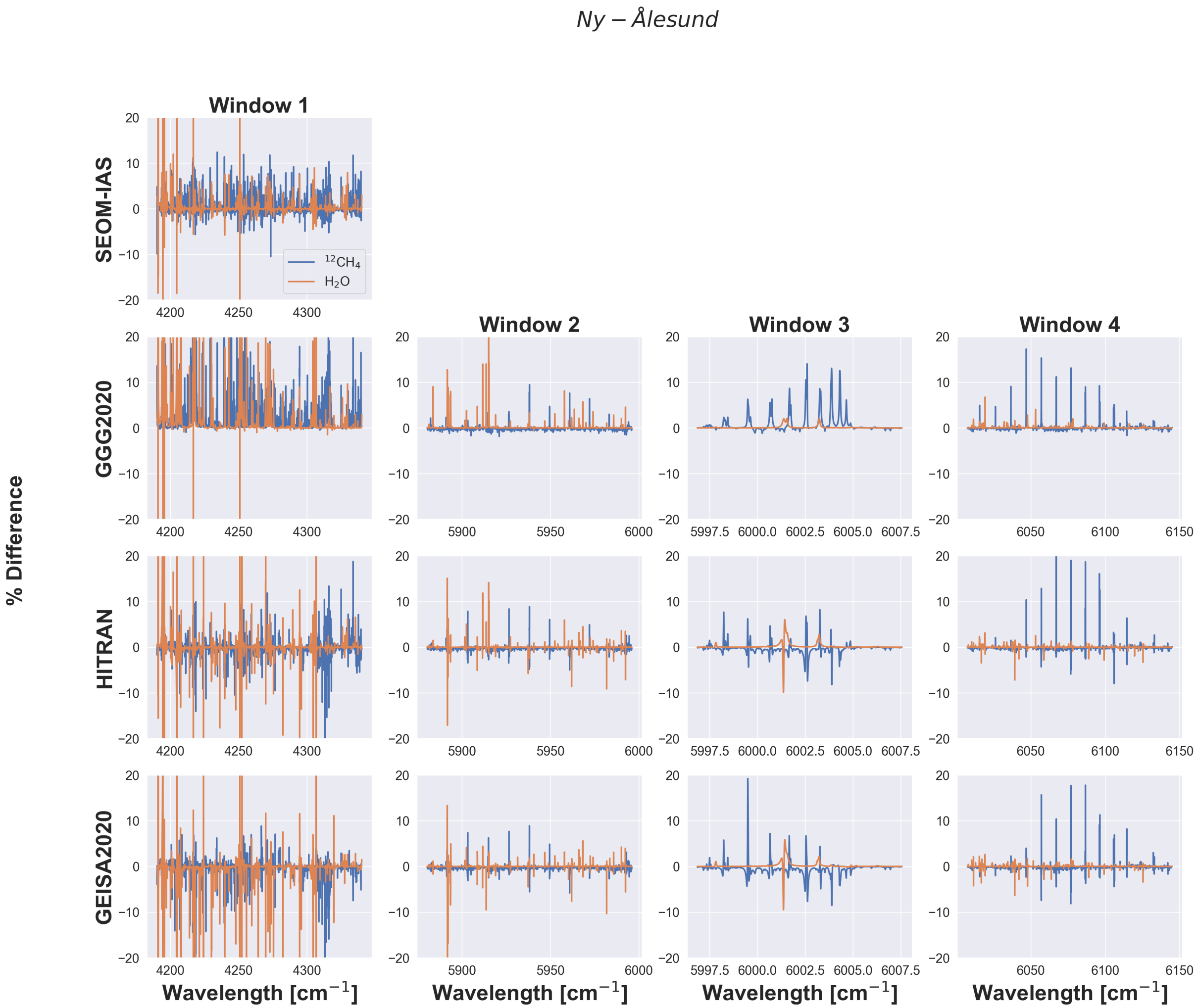


Figure 3. Percent difference in calculated transmission for $^{12}CH_4$ and H_2O from each spectroscopic database for each window w.r.t. GGG2014. Rows indicate spectroscopic database and columns indicate the windows. The y-axis values have been limited to 20% to avoid the plots being dominated by large excessive noise values, note however, especially in window 1, the values are sometimes in excess of 20%

The results shown in Fig. 3 suggest a number of different conclusions. Firstly the impact of differences in the spectroscopic parameters of water vapour are highly significant in window 1, more so than differences in $^{12}CH_4$, with each spectroscopic database showing differences >20% at numerous wavelengths. Further, each spectroscopic database shows significant disagreement as to where the differences occur, suggesting large differences in the treatment of water vapour parameters in window 1 for each spectroscopic database. Therefore the poorer fit quality shown in Fig. 1 for window 1, is likely driven by water vapour uncertainty as opposed to methane. For window 2, we note again that water vapour seems to have the largest uncertainties (not to the degree of window 1). The main points of disagreement in general line up with the largest deviations in Fig. 1. For windows 3 and 4, water vapour differences have less impact, with the majority of the differences attributed to $^{12}CH_4$, which still have a lower magnitude than window 1. The large uncertainty in window 3 at 6001 cm^{-1} (characteristic bump shape), seems to be caused by water vapour uncertainty in HITRAN and GEISA2020. The main conclusion from this assessment, is that when considering uncertainty in $^{12}CH_4$ between windows and spectroscopic databases, uncertainty in water vapour should be considered at the same time.

3.2 Quantification of variance between windows and databases

The entire time series available for this study for each TCCON site are shown in Figs. 4, 5, 6 and 7. XAir, a quantity normally retrieved with TCCON is shown as an additional quality indicator for GGG2014, GGG2020, HITRAN and GEISA2020 (SEOM-IAS does not have spectral coverage in the TCCON XAir retrieval window) with variations between 0.96 and 1.04 assumed as good quality. Qualitative inspection of these figures shows scatter between all windows for each database, further the HITRAN, ~~GESIA~~ GEISA and SEOM-IAS databases show significant positive bias w.r.t. the standard deviation of the reference TCCON retrieval, indicated by the dashed black lines. Quantitative metrics for these figures are shown in Fig 8. We also find the XAir values indicate quality retrievals, with only a small number falling outside the acceptable range.

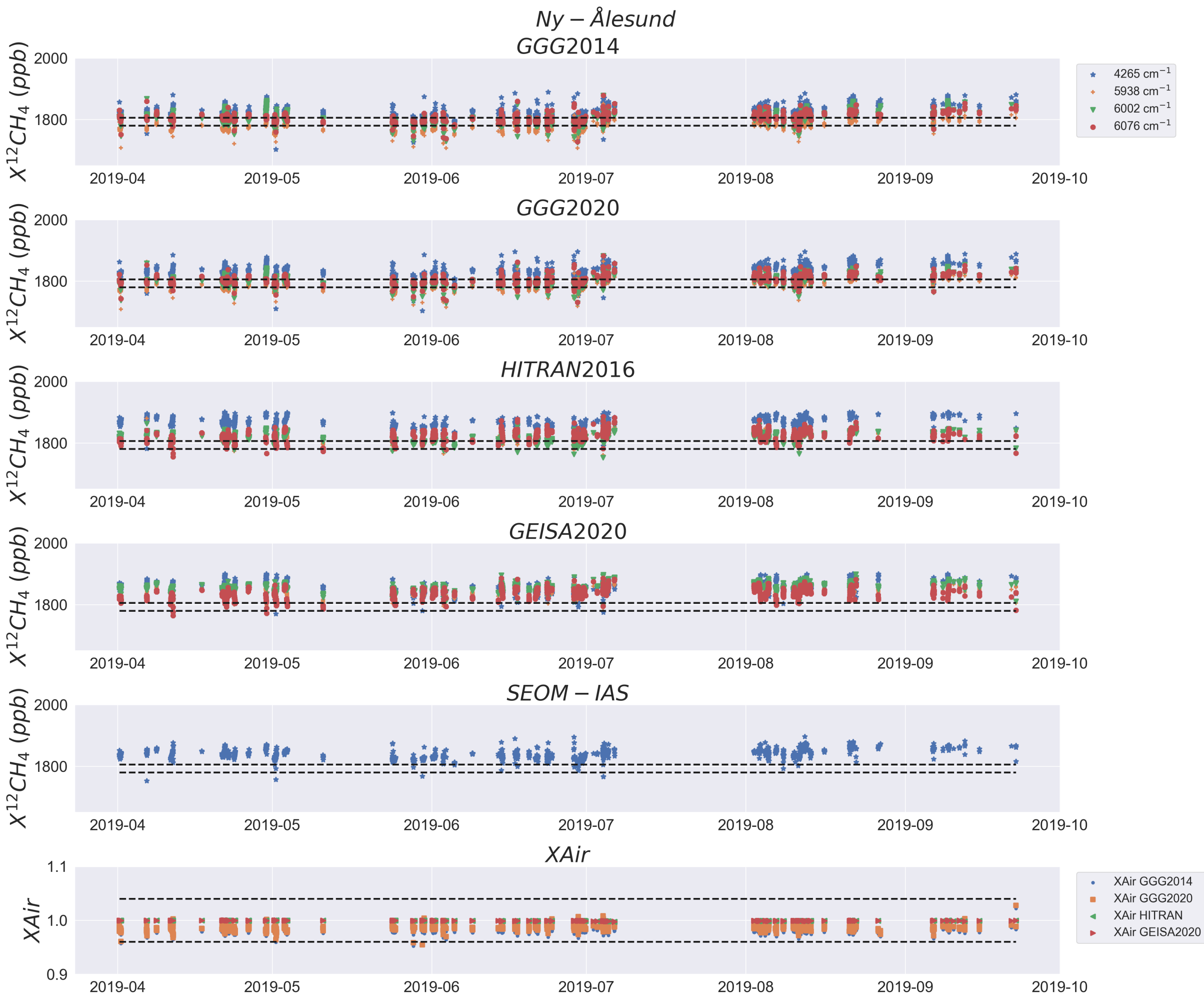


Figure 4. Retrieval time series for $^{12}CH_4$ DMFs from the Ny-Ålesund site. Each panel indicates retrievals from each spectral window (indicated in the legend) from a specific spectroscopic database, indicated in the panel title. Blue stars show retrievals from band 1, yellow pluses are band 2, green triangles are band 3 and red circles are band 4. The standard deviation about the reference TCCON retrievals are indicated by the horizontal dashed lines. The bottom panel indicates the retrieved XAir DMF as a quality indicator for the retrievals, with the dashed lines indicating the standard range of acceptable XAir values.

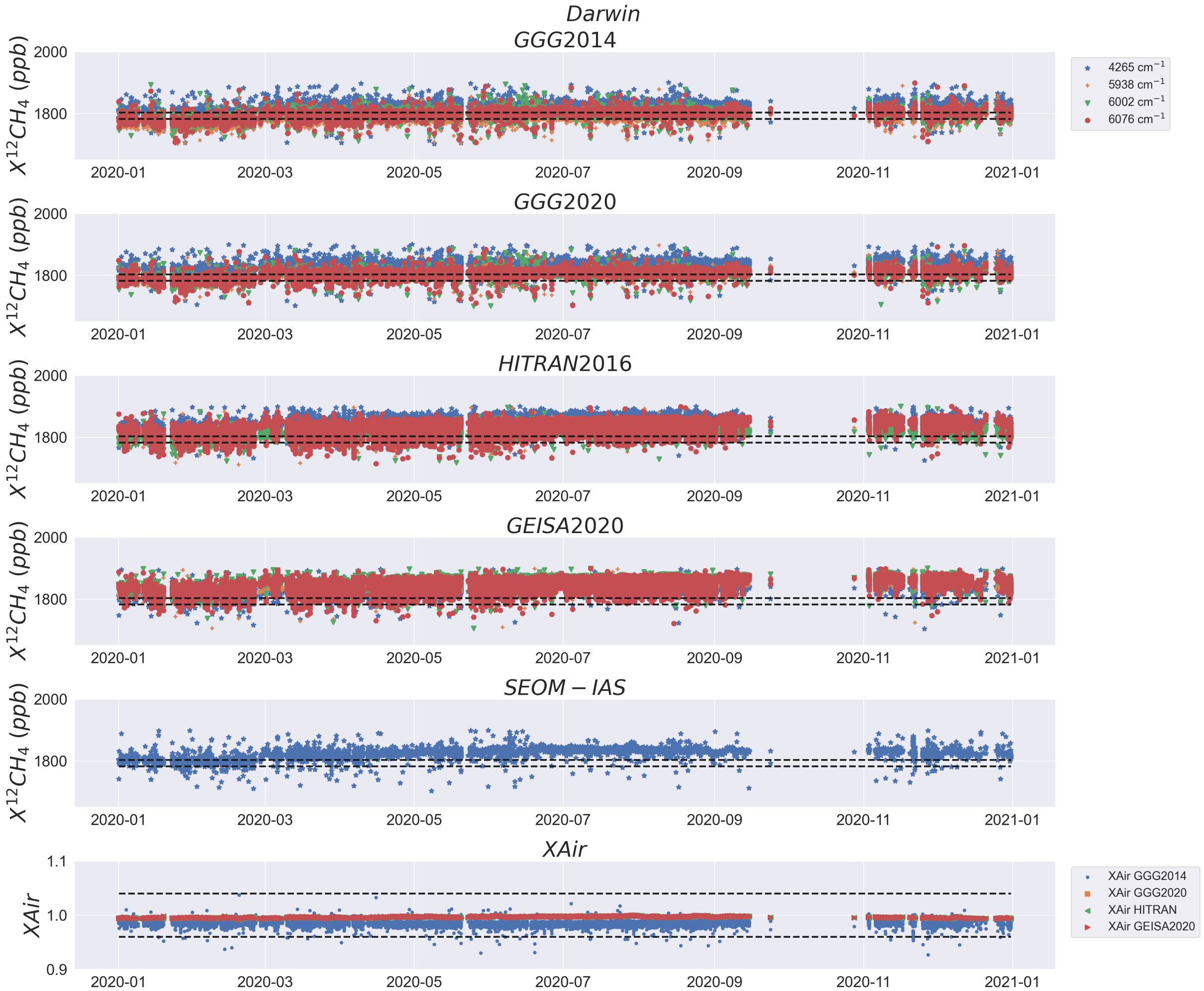


Figure 5. As Fig. 4, but for retrievals from the Darwin TCCON.

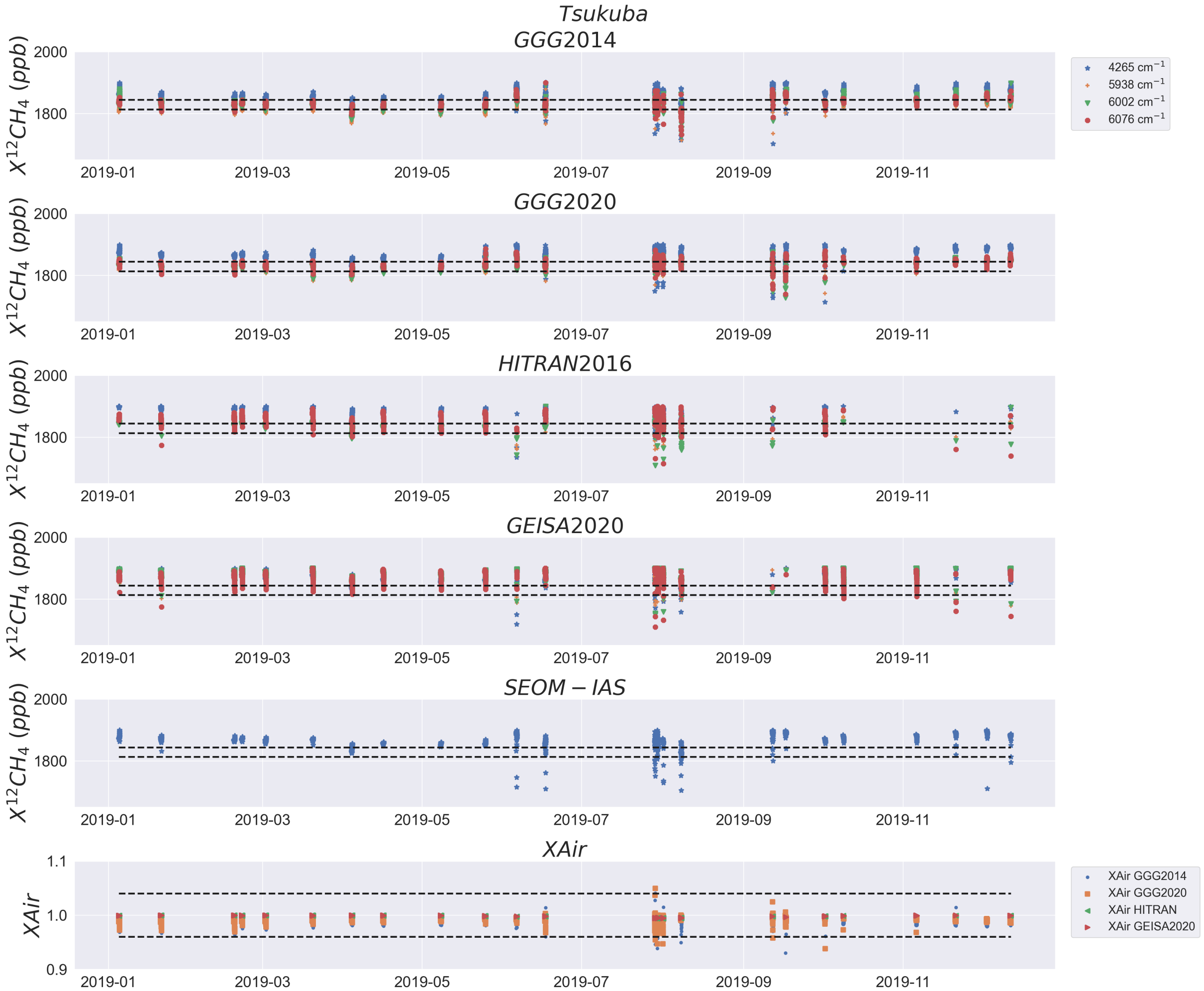


Figure 6. As Fig. 4, but for retrievals from the Tsukuba TCCON.

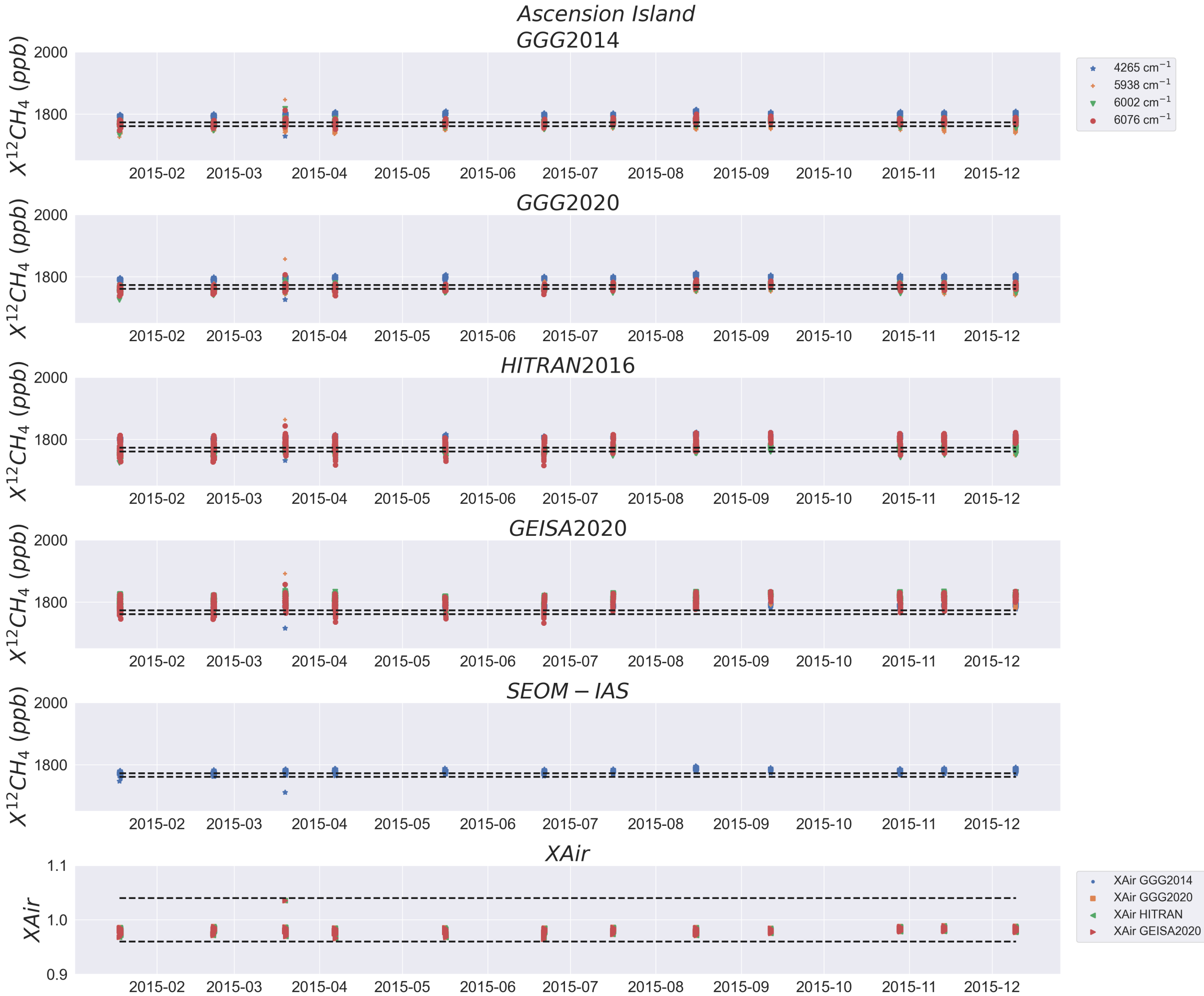


Figure 7. As Fig. 4, but for retrievals from the Ascension Island TCCON.

The metrics used in Fig 8 indicate ~~the the~~ that the bias for $^{12}CH_4$ retrievals for each database and window w.r.t the reference retrieval (Norm Abs Mean), and the presence of any large deviations (RMSE). These metrics are normalised by the retrieval uncertainty of the reference retrievals a weighted average of windows 2, 3 and 4 from GGG2014, thus we assume any biases with values greater than 1 cannot be attributed to uncertainty and are therefore real.

Firstly we analyse the results from Ny-Ålesund, which due to the constant nature of the atmospheric conditions can be considered as a baseline. For window 1, both the NRMSE and Norm Abs Mean values for all of the databases indicate values greater than 1, thus suggesting there are still significant variations in the treatment of spectroscopic parameters in window 1. The HITRAN and GEISA databases show bias deviations twice that of GGG2014, however these values do not indicate any one database is more accurate than the other, but either large differences in spectroscopic parameters or differences in sensitivity to local conditions. Windows 2 & 3 do not show any notable biases apart from the GEISA database which generally shows the largest deviations across all of the windows (except window 1). In window 4, both HITRAN and GEISA show notable deviation from the reference retrievals, which is a surprising result given this window is popular in satellite retrievals of methane (Yoshida et al., 2013). We note the NRMSE and Norm Abs Mean values are similar in the majority of cases, indicating that there is an underlying bias between the database retrievals as opposed to large spikes of differences. Considering the bias deviations across the windows, GEISA is the only example to exceed values of 1 across all windows, with window 3 showing the largest deviation from the reference value.

Secondly considering the dataset from Darwin, the magnitude of the NRMSE and Norm Abs Mean values are typically lower than the equivalents in the Ny-Ålesund dataset. The relative differences between the NRMSE and Norm Abs Mean values between the databases are the same as those shown in the Ny-Ålesund dataset, i.e. GGG2014 shows the lowest differences and GEISA shows the largest, apart from ~~in~~ window 1 in which case it is HITRAN. Investigating each window in turn, only HITRAN shows a notable deviation from the standard retrieval in window 1 with GGG2020 and GEISA not indicating a significant deviation above the standard noise level (only 0.02 and 0.07 above 1 respectively). For windows 2 & 3, only the GEISA database shows a significant bias with respect to the standard, as with Ny-Ålesund site. Again in window 4 only the HITRAN and GEISA databases show notable deviation from the standard, again suggesting the spectroscopic parameters in window 4 still have significant uncertainty w.r.t windows 2 & 3. The implication of the Darwin results w.r.t those from Ny-Ålesund are that either or both the differences in the instrument setup and the local conditions impact inter-window/spectroscopic database biases. Note for the Darwin retrievals there is an inconsistency at the beginning of March 2020, where there is a small 'bump' in the magnitudes, and lower magnitude results are not apparent. This effect only appears at the start of March, and the typical retrieval variability is quickly restored. Quality control indicators (such as XAir) do not indicate any problems with the results in this period, and the reason for this inconsistency remains unclear.

The results for the Tsukuba retrievals are very similar to those shown for Ny-Ålesund, with GGG2014 not showing any significant differences except in window 1, as with the other sites HITRAN shows deviation in windows 1 and 4, and GEISA showing the largest differences apart from in window 1. However, the main difference is with the GGG2020 database, as with the other TCCON sites the normalised absolute mean shows deviation in windows 1 and 4, however the NRMSE indicates significant differences in all windows, suggesting there are a small number of retrieval cases that have large biases w.r.t the standard values. This behaviour is not replicated in the other TCCON sites.

Finally, all results from the Ascension Island measurements indicate no deviations of any significance, contrasting with the results from all other ~~sides~~sites. We note the standard deviation about the reference TCCON retrievals in Fig 7 is smaller than any of the other TCCON sites. This suggests constant retrievals in methane over the course of the year at Ascension Island, and therefore limited opportunity for biases to form.

The results in Fig 8 clearly indicate that in the cases where deviations exist, they are reflected in all of the TCCON sites (when significant), implying that despite the fit differences shown in Fig 2 these biases cannot (purely) be attributed to errors in the TCCON instruments, but given the consistency of the deviations we can attribute these differences to spectroscopic parameters. Figure 8 indicates that there are significant differences between SEOM-IAS, GEISA and HITRAN databases w.r.t. the GGG databases, which show less deviation. This is not surprising since the reference values are based on GGG2014, and GGG2020 is built upon GGG2014, however this is not the case in window 1 where larger deviations are observed. This suggests that knowledge of spectroscopic parameters in window 1 is still not as settled as the other windows which have been routinely used in TCCON. It is difficult to assess all of the differences between the databases, due to the range of parameters used; there are some papers which describe the sources of the spectral lines for each of the databases (Brown et al., 2013; Jacquinet-Husson et al., 2016), but specifics are limited due to the size of the databases. Complexity is added by the fact that several of these databases state that data is drawn from the same sources (Albert et al., 2009; Nikitin et al., 2015, 2017), however these papers go on to say that not all of the lines from these studies are implemented based on in house assessments of fit quality. The implication being that it is challenging to specifically identify where spectroscopic parameter differences occur between the databases.

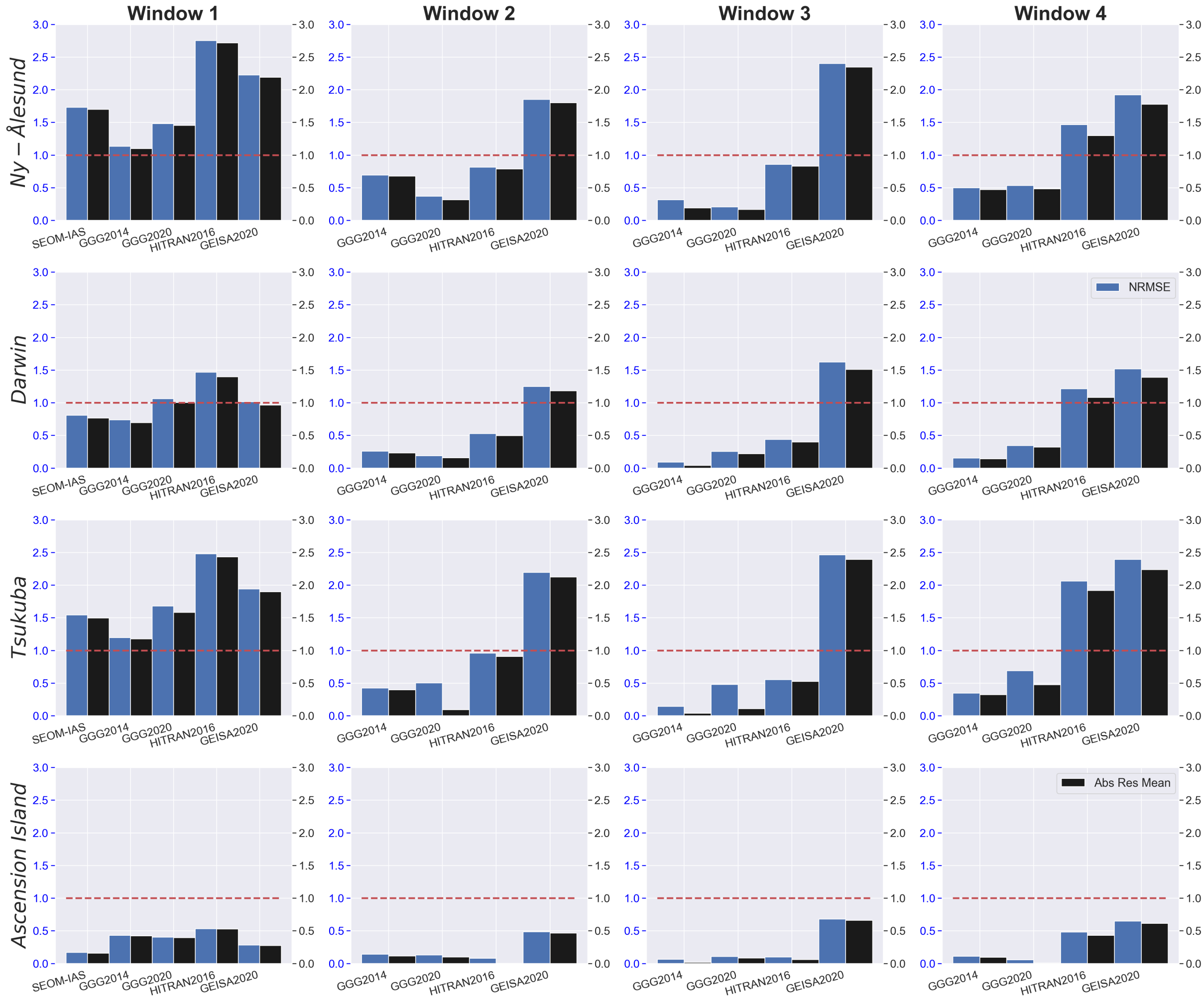


Figure 8. Bar plot indicating NRMSE and Normalised Absolute mean residual difference values for $^{12}CH_4$ retrievals from each TCCON site, each window and each spectroscopic database under consideration in this study with respect to the original TCCON methane retrieval window. Each row shows data from each TCCON site, as indicated by the y-axis titles, and each column shows results from each window, as indicated by the column title. Each subplot shows the NRMSE (the blue bars, magnitude shown by the left-hand y-axis) and Abs Res Mean (the black bars, indicated by the right-hand y-axis) values for each spectroscopic database as indicated by the x-axis. The horizontal red-dashed lines indicated the magnitude of 1, the value where we assume the bias values to be significant.

For $^{13}CH_4$ DMFs there is no obvious reference value available, since $^{13}CH_4$ is not typically retrieved from TCCON. We therefore chose to use GGG2014 window 1 as a reference in order to investigate window deviations. We found that apart from SEOM-IAS window 1 which showed deviation below the noise level from every TCCON site, every other case showed notable levels of deviation ranging from 1.5-5. Here we cannot attribute these disagreements purely to spectroscopic differences since $^{13}CH_4$ retrievals will be subject to high noise levels.

3.3 Impact of local condition changes on variance between windows and databases

~~The TCCON sites used in this study were picked to have a wide range of conditions, with Ny-Ålesund capturing spectra in largely unvarying conditions with high SZA and low water vapour, while Ascension Island is similar in unvarying conditions although with higher background water vapour conditions and lower SZAs. This is contrasted by Darwin and Tsukuba which capture spectra under a wide range of SZAs and highly variable water vapour conditions (see Table 2).~~ It has been shown (Wunch et al., 2011) that the variability of local conditions can have an impact on the accuracy of TCCON retrievals (through the ~~apriori~~ a priori data). We therefore investigate in this section if varying local conditions (specifically, water vapour, SZA and temperature) affect each window in each spectroscopic database differently. Wunch et al. (2011) identified a non-linear relationship between SZA and retrieval anomalies, we therefore fit the normalised residual DMF values with a second order model, to expose any potential non-linearity.

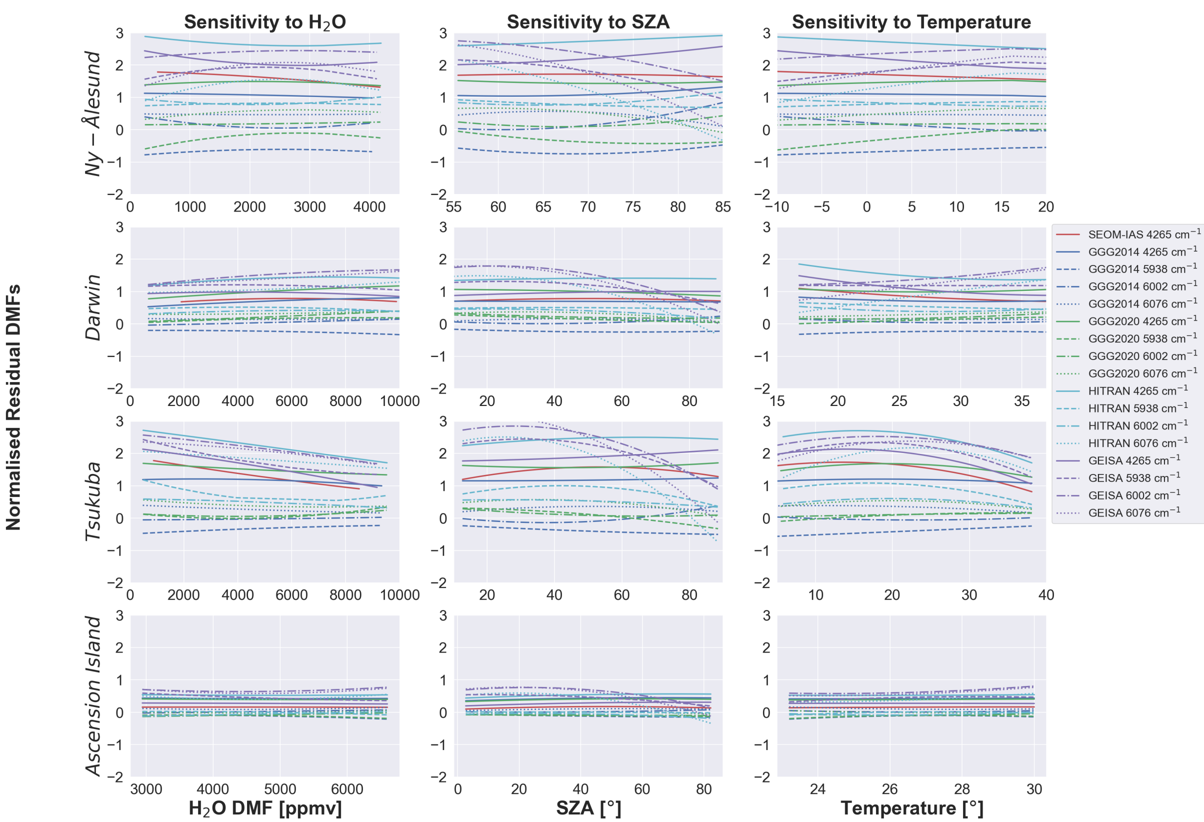


Figure 9. The sensitivity of water vapour, SZA and temperature variations on retrieved $^{12}CH_4$ DMFs from each TCCON site, spectroscopic database and windows. Each subplot shows the ~~linear~~ second order regression fit of the normalised residual between the retrieved window/spectroscopic database DMFs and the TCCON reference DMFs, with the fit window and spectroscopic database indicated by the legend. Each row of the figure shows data from each TCCON site as indicated in the y-axis, and each column shows the sensitivity to a specific condition as shown in the title.

~~The qualitative distributions indicated in Fig. 9 are~~ Figure 9 qualitatively describes the sensitivity of each TCCON site/database/window to variations in local conditions. For Ny-Ålesund (row 1), there is a mixture of non-linear and linear sensitivities to variations in water vapour and SZA. Windows 2, 3 and 4 for GEISA2020 indicate particularly significant non-linear sensitivities to SZA variations. Sensitivities to temperature variation are generally linear, although some indications of slight non-linear behaviour are apparent (GEISA2020). For Ny-Ålesund there are some cases where little sensitivity is observed, e.g. HITRAN, suggesting a wide range of responses in the databases/windows. In contrast to Ny-Ålesund, Darwin (row 2) shows limited sensitivity to local condition variations, with low magnitude linear gradients observed for most cases. There are some exceptions, notably HITRAN window 3 and GEISA2020 windows 3 and 4 in relation to SZA variations, where significant non-linear behaviour is observed. Tsukuba (row 3) again shows significantly different behaviour, with almost all databases/windows showing significant linear or non-linear sensitivity. Window 1 for SEOM-IAS, GGG2020, HITRAN and GEISA indicate significant negative linear relationships, with all other cases showing a range of sensitivity. For variations in SZA, as with Ny-Ålesund and Darwin, HITRAN window 3 and GEISA2020 windows 3 and 4 suggest strong non-linear sensitivity to variations in SZA. Most of the other windows/databases indicate some linear/non-linear sensitivity, but not to the same degree as HITRAN window 3 and GEISA2020 windows 3 and 4. Temperature variations for Tsukuba indicate significant non-linear sensitivity for window 1 in most cases (except GGG2014), and in general show different results from those shown in Ny-Ålesund and Darwin. Finally for Ascension Island, we note almost no sensitivity to any local condition variation, except for HITRAN window 3 and GEISA2020 windows 3 and 4 with SZA variations, which have shown sensitivity in all cases.

The qualitative assessment above is explored quantitatively in more detail in Figs 10, 11, and 12. Where the Pearson's correlation coefficients and the regression coefficients for a second order fit are shown. Note that the y-axis scales are unified for all sites and windows for each variable considered, to allow for direct comparison. In the following analysis, we assume the presence of a substantial linear correlation when r values > 0.5 are shown. Further, we assume any deviation from zero of the regression statistics to be the presence of at least minor sensitivity. First considering the impact of SZA variations in window 1 across all sites~~, there is no clear pattern in the correlations. For example,~~ . There is no indication of a substantial linear correlation, with no database at any of the sites showing an r value > 0.5. The regression statistics for Darwin and Ascension Island do not suggest any non-linear or linear relationships, while Tsukuba indicates the presence of a non-linear relationship for SEOM-IAS and HITRAN. All three sites do indicate the presence a constant bias (although minor in the case of Ascension Island). The results from Ny-Ålesund indicate more significant sensitivity to SZA variations. SEOM-IAS shows a slight linear relationship, while GGG2014~~at Ny-Ålesund and Ascension Island shows a weak correlation of ∼0.25, yet no correlation at either Darwin and Tsukuba. This pattern is repeated for all databases except GEISA2020~~, ~~which shows a weak correlation across all TCCON sites~~GGG2020 and GEISA2020 show second and first order sensitivity, in contrast HITRAN shows only minor sensitivity. For window 2 we see weak to strong linear correlation in all databases across all sites, except for a few cases (e.g. HITRAN at Darwin and most cases at Ascension Island). The regression statistics indicate minor sensitivity for GGG2014, GGG2020 and HITRAN at most sites, except for Tsukuba where HITRAN is significant. Ny-Ålesund shows significant non-linear regression statistics for all databases, indicating significant sensitivity to SZA variations, and given that Ny-Ålesund operates at the highest SZAs of all of the TCCON sites under consideration, this is logical. Window 3 shows

~~high~~ notable levels of correlation to SZA variations w.r.t. all sites, with Ny-Ålesund, Tsukuba and Ascension Island showing particularly ~~notable correlations~~ large correlations (>0.5), especially in the GGG2014 and GEISA2020 databases. However, Darwin and Ascension Island show only very minor sensitivty to SZA variations, except for GEISA2020 at these sites. Tsukuba and again Ny-Ålesund shows much higher sensitivity for all databases (except GGG2020 and HITRAN for Tsukuba), with the regression statistics showing 2nd and 1st order coefficient values much larger than any other case. Window 4 shows large ~~correlations for all databases~~ linear correlations for almost all of the databases, in the Ny-Ålesund retrievals especially~~, but large negative correlations across all sites with the HITRAN and GEISA~~ . Generally for Darwin, Tsukuba and Ascension island, little to no sensitivity to SZA variations are observed for GGG2014 and GGG2020, while HITRAN and GEISA2020 all indicate more significant sensitivity, comparable to Ny-Ålesund in window 1. In contrast, Ny-Ålesund shows a much greater sensitivity to SZA variations for all spectroscopic databases. Overall the results in Fig 10 indicate ~~the retrieval biases w.r.t the reference TCCON retrievals are sensitive to SZA variations.~~ limited to no sensitivity to SZA variations for the Darwin, Tsukuba and Ascension Island sites for the GGG2014 and GGG2020 databases, while HITRAN and GEISA2020 do indicate some sensitivity, especially in windows 3 and 4. The Ny-Ålesund site~~shows the largest sensitivity of~~ , however, shows significant sensitivity for all of the spectroscopic databases, across all of the ~~sites considered in this study. This~~ windows. With no clear 'winner' or 'loser' in terms of windows and spectroscopic databases, with all statistics showing similar results. These results could be explained by the fact that Ny-Ålesund operates at higher SZA angles than any of the other TCCON sites, meaning the retrieval path length will be longer, potentially allowing for more errors to ~~creep~~ propagate into the retrievals. ~~However, the~~ The results from Ascension Island, which operates at lower SZAs than any of the other TCCON sites considered in this study ~~also indicates large correlations, suggesting further complexity~~indicates the lowest magnitude sensitivities, adding weight to this argument.

The sensitivity of each window to water vapour variation is explored in Fig 11. ~~For~~ Beginning with window 1, ~~there is weak sensitivity to water vapour variations at the~~ only the SEOM-IAS, HITRAN and GEISA2020 databases at the Tsukuba site indicate the presence of a significant linear correlation. For these sites we note that the second order regression statistics indicate only very minor sensitivity, and although there is more evidence of stronger first order sensitivity, the lack of a linear correlation suggests water vapour variation only has a minor impact. Except at Tsukuba which is an interesting results, given both Tsukuba and Darwin both have large variations in background water vapour. Ny-Ålesund ~~and the Darwin sites, with opposite databases showing greater sensitivity. The Tsukuba site shows large negative correlations for the HITRAN, GEISA and SEOM-IAS databases, while Ascension Island shows very little correlation across all databases. Window~~ however, indicates the presence of strong non-linear relationships for GGG2020, HITRAN and GEISA2020, contrasting the results from the other sites. Window 2 ~~shows large correlation magnitudes, with the Ny-Ålesund and Tsukuba both showing correlations >0.5 in several cases. In general, all databases in window 2~~ shows ~~correlation values >0.25~~ little linear correlation for most of the cases, except for ~~a small number of cases, most notably at the Darwin site. Different patterns are observed with windows 3, where Ny-Ålesund shows some significant negative correlations , while Ascension Island indicates sensitivity across all databases (except GEISA). Darwin and Tsukuba show opposite behaviours with large positive correlations at GGG2014 and~~ GGG2020 at ~~Darwin, but large negative biases with the HITRAN and GEISA databases at Tsukuba. For window 4,~~

~~notable correlations are observed for GGG2020, HITRAN and GEISA at~~ Ny-Ålesund, ~~but in general~~ and all sites at Tsukuba. Ny-Ålesund indicates the presence of significant non-linear correlations for all spectroscopic databases in contrast to Darwin which shows only very minor sensitivity to water vapour variations. GGG2014 at the Tsukuba site presents a very minor linear sensitivity, while HITRAN and GEISA2020 both show the presence of a significant non-linear regression. Finally for window 2, Ascension Island shows similar results for GGG2014, GGG2020 and HITRAN, namely a minor non-linear relationship, while GEISA2020 shows the presence of a more significant non-linear relationship. With window 3, there are no ~~other correlations of note across the other TCCON sites(except for~~ cases of significant linear correlation; we note no or very slight (HITRAN and GEISA2020 Darwin) non-linear relationships at Darwin, Tsukuba and Ascension Island, except for GEISA2020 at Ascension Island. The results from Ny-Ålesund contrast the other sites, by indicating strong non-linear relationships for all spectroscopic databases, except GGG2020. Window 4 shows similar results to window 3, with almost no cases indicating significant linear correlation (except GGG2014 at Tsukuba). ~~These results are curious, we know Darwin , Tsukuba and Ascension Island have much higher background water vapour levels and variability ~x3 than~~ Darwin and Tsukuba show only very slight non-linear sensitivity to water vapour variation, while Ny-Ålesund and Ascension Island both indicate significant sensitivity across all databases (aside from GGG2014 at Ny-Ålesund~~. Yet, Ny-Ålesund shows larger correlation in some cases than any of the other sites with high water vapour concentrations. Tsukuba has a~~). The conclusions from this analysis suggests locally varying conditions are key in determining the impact of water vapour variations. It is interesting that the site with the lowest magnitude background water vapour ~~content less than that of Darwin, but exhibits variability ~x2 that of Darwin , and for Tsukuba there is much greater indication of bias~~ (Ny-Ålesund) is the most affected by water vapour variation, as opposed to Darwin which has the highest background and variation, which shows little sensitivity to water vapour variation. The key difference between Ny-Ålesund and the other sites is the SZA at which measurements are taken, meaning TCCON measurements taken at high SZA will be more sensitive to other varying conditions. ~~Thus suggesting that the variability of water vapour in the atmosphere is more significant for retrieval biases, as opposed to high background levels.~~

The results for retrieval bias sensitivity to variations in temperature are shown in Fig 12. ~~Comparisons of Figs 11 and 12 show similar results in both cases, for example,~~ For window 1~~at the Ny-Ålesund site shows almost identical results. This suggests for the~~, only the Tsukuba site shows significant linear correlation with temperature variation, for the SEOM-IAS, HITRAN and GEISA2020 databases. Ny-Ålesund ~~site, water vapour and temperature variability occur on similar time scales, this result is expected as the atmosphere can hold more water with rising temperatures. Similar patterns are apparent at the Tsukuba site with some notable exceptions, although these exceptions occur in cases with low bias sensitivity (e. g. GGG2014 window 3), meaning these are unimportant. However, this similarity is not as apparent at the Darwin site, with~~ shows limited sensitivity across all spectroscopic databases, contrasting with the other TCCON sites, all of which indicate some sensitivity. Darwin and Tsukuba show the presence of a non-linear relationship across all spectroscopic databases, with Ascension Island showing a non-linear relationship in most cases. For window 2, Ny-Ålesund shows very limited non-linearity, and although strong linear correlations are observed, only linear coefficients of minor magnitude are observed, indicating low sensitivity. Darwin shows little to no sensitivity across all cases, while Tsukuba suggests a strong linear correlation for GGG2014, but almost no sensitivity, but both HITRAN and GEISA2020 suggest strong non-linear relationships. Contrasting the other sites, Ascension island shows almost no linear correlation, but the presence of significant non-linear relationships for all spectroscopic databases. For window 3~~showing higher retrieval bias sensitivity to water vapour than to temperature variation. Indeed, Darwin~~ , Ny-Ålesund shows only minor sensitivity to temperature variations, while Darwin indicates the presence of non-linear relationships similar to window 1. Tsukuba shows similar results to window 2, such that GGG2014 does not indicate any ~~cases with retrieval bias correlation >0.25, (unlike any of the other sites). The key exception to this pattern are the results from Ascension Island, where the values for water vapour and temperature sensitivity are significantly different~~sensitivity, while HITRAN and GEISA2020 show stronger non-linear relationships. Again Ascension island is the outlier, showing strong non-linear sensitivity for all spectroscopic databases. Window 4 is also different from the other windows, with the results from Ny-Ålesund generally showing more sensitivity to temperature variations than the other windows, especially GEISA2020. For Darwin, GGG2014 and GGG2020 show very little sensitivity, while HITRAN and GEISA2020 show significant non-linear relationships. Tsukuba shows similar results to windows 2 and 3, where HITRAN and GEISA2020 indicate strong non-linear sensitivities, while GGG2014 is largely invariant to temperature variations. Ascension Island shows similar results to windows 2 and 3, where strong non-linear relationships are observed for all spectroscopic databases. In summary, variations in temperature will impact inter-window and inter-database biases, the impact depends significantly on the local conditions as well as the window and database in question. We note ~~in this case that the temperature variation at Ascension Island (Table 2) is very small, which could explain these differences~~that Ascension Island shows the most significant impact in these results, however this is possiblly a biased result due to the fact that Ascension Island shows very little temperature variation over the measurement dataset. The other sites, which capture measurements over a much wider range of temperatures are therefore more reliable.

In general, there is no clear case of one window ~~, database or TCCON site~~ or database showing clear sensitivity over and above than any of the others ~~in all cases, meaning one site or database is not especially sensitivethan the others~~(although Ascension Island is typically less sensitive). However there are clear indications of sensitivity to variations in the local conditions which vary between window, database and TCCON site, in some cases very strong correlations ~~. In general, the pattern is that variability in the local conditions causes window and database biases, rather than extreme conditions by themselves~~and sensitivities. For example, Ascension Island has some of the least varying conditions of all of the sites, and ~~also shows the most constant sensitivityto varying conditions. While Tsukuba has some of the most variable conditions and also indicates some of the most variability when assessing water vapour variability. This assessment is not perfect, since~~ the results from this study indicate the least bias sensitivity. While Ny-Ålesund ~~also shows significant dependence on local conditions, while having~~ has less variability than Darwin or Tsukuba~~. However, Ny-Ålesund spectra are captured at high SZA meaning lower SNR and more susceptibility to interfering elements~~, but takes measurements under more challenging conditions, and also shows more sensitivity than either of these sites.

We ~~also~~ note when calculating the Pearson's correlation ~~coeffecient~~ coefficient for GGG2020 values for window 1 at the Tsukuba site, large 'p' values were found, indicating these results are not statistically significant, and therefore should be ignored. The p-test was applied to all other window and spectroscopic database combinations, all showed significance w.r.t the p-test i.e. «< 0.05.

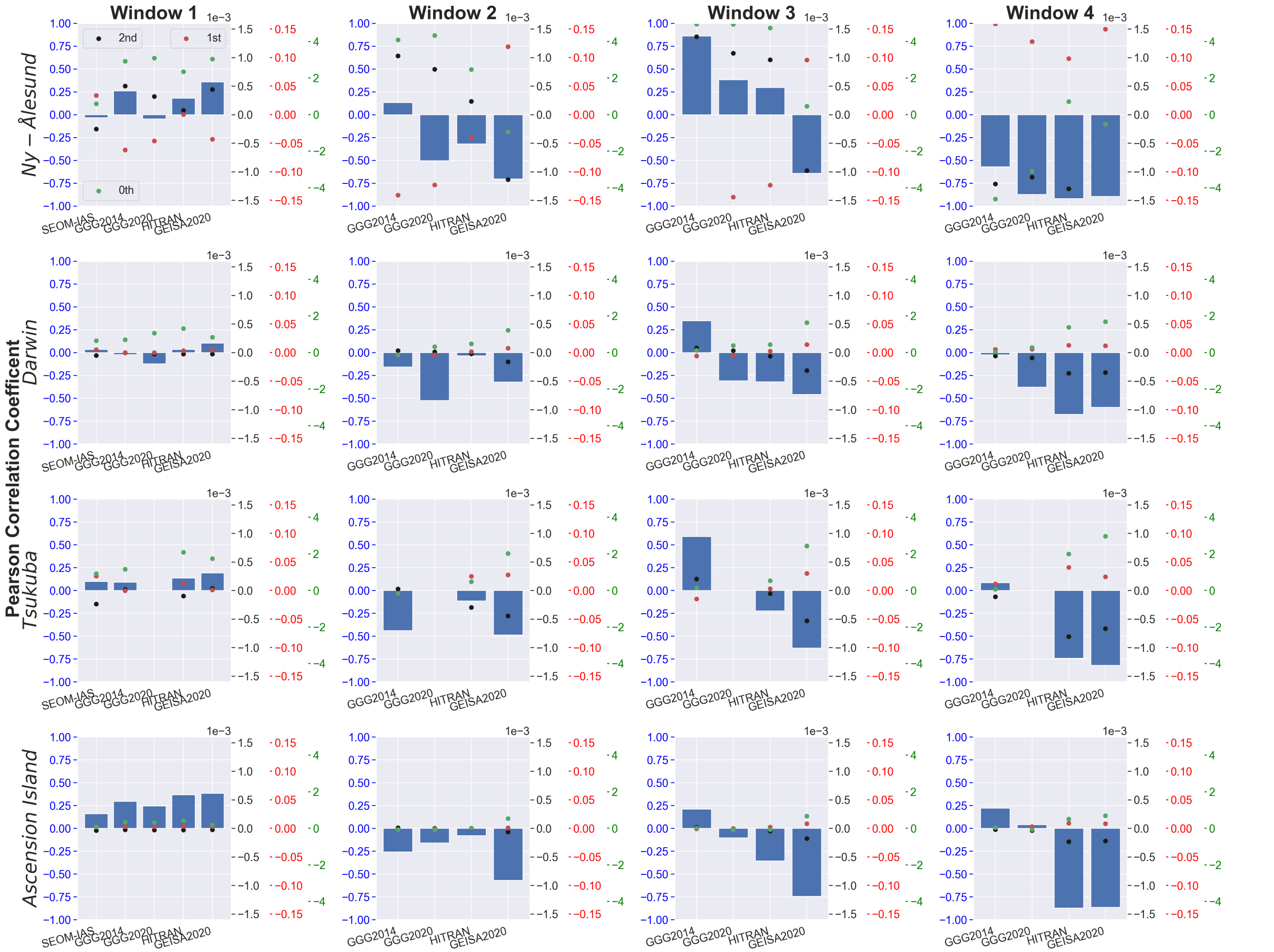


Figure 10. Bar and scatter chart indicating the statistics behind the sensitivities to SZA variations shown in Fig 9. Each row indicates the results from a TCCON site, as indicated by the y-axis label, each column shows the results from a particular window as shown by the column title. The blue bar plot show the Pearson correlation ~~coeffiecent~~coefficient, with the left-hand blue y-axis values the appropriate scale. The black dots ~~showing~~ show the ~~linear gradient of~~ second order coefficient from the ~~linear fits~~ regression from the fits shown Fig 9, with the right-hand black y-axis values as the scale. The red dots show the first order coefficient, corresponding to the right-hand red y-axis, and the green dots show the constant values corresponding to the right-hand green y-axis. Note all values for GGG2020 at the Tsukuba site have been removed due p-test failure.

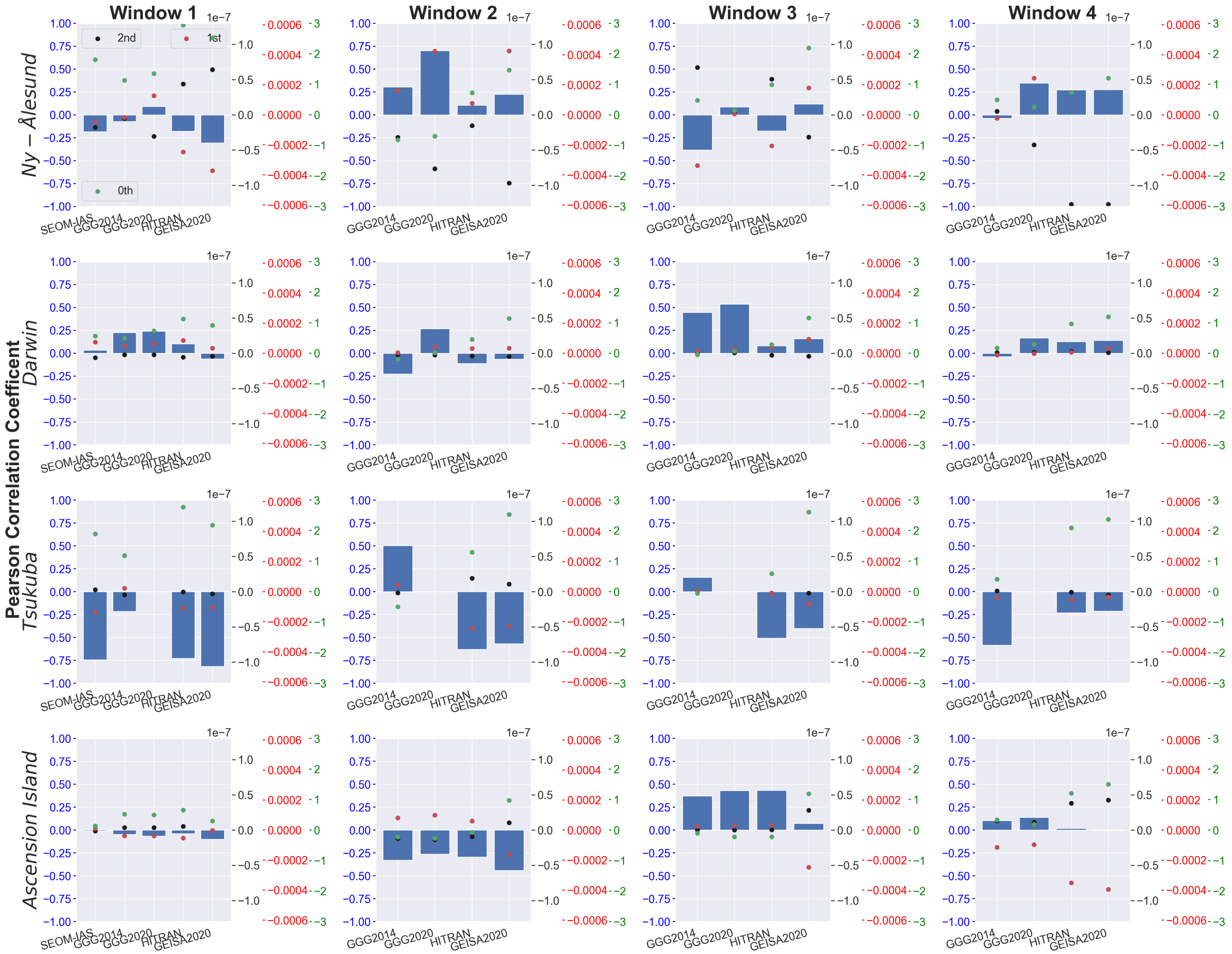


Figure 11. As Fig 10, but showing the sensitivities to water vapour variations.

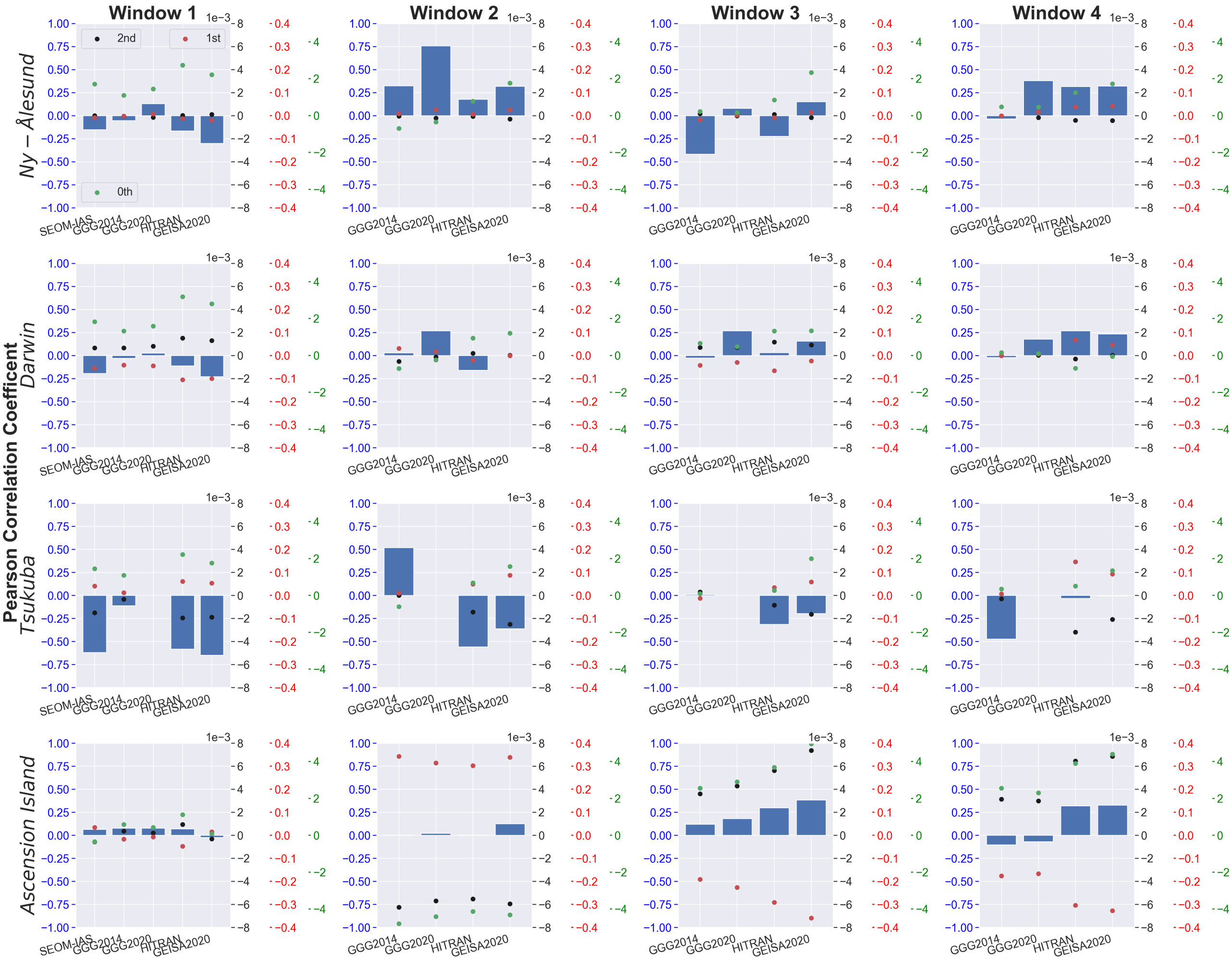


Figure 12. As Fig 10, but showing the sensitivities to temperature variations.

A similar analysis for retrieval bias sensitivities for $^{13}CH_4$ indicated high levels of sensitivity to SZA variation, especially those retrievals from Ny-Ålesund where SZAs are high. There are some windows that indicate no correlation, but the majority had values greater than 0.3. W.r.t water vapour and temperature variation, these results are mixed with different windows and databases at different sites indicating different results. However there is a general trend of sensitivity to water vapour and temperature variations, with only a small number of cases indicating no correlation. These results suggest $^{13}CH_4$ retrievals are more sensitive to changing conditions than $^{12}CH_4$, which is the expected result.

3.4 Calculation of δ^{13}C values

The calculation of the δ^{13}C values (Eq 1), can give some insight into the accuracy of $^{13}CH_4$ retrievals from TCCON, as well as the impact of local condition variations on these retrievals. δ^{13}C is calculated for all TCCON sites using all combinations of windows from all databases~~in Table ??~~, using averaged $^{12}CH_4$ and $^{13}CH_4$ for the whole time series available for each TCCON site. There are two factors to look for in the ~~calculation~~ analysis of $\delta^{13}C$, firstly the bias w.r.t. the accepted atmospheric average of -47‰ and the consistency of the calculated values across databases and windows.

Table 4. Averaged values of $\delta^{13}C$ from all TCCON sites for all possible combinations of $^{12}CH_4$ ~~and~~ with window 1 of $^{13}CH_4$ ~~window combinations~~ for each spectral database.

~~Windows $^{12}CH_4$ & $^{13}CH_4$~~ Site	Database	Windows 1 & 1	Windows 2 & 1
Ny-Ålesund	GGG2014: -102‰ GGG2020: -90.6‰ HITRAN: 8.98‰ GEISA: -66.4‰ ~~SEOM: -91.3‰~~ GGG2014 GGG2020 HITRAN GEISA SEOM	GGG2014: -77.3‰ GGG2020: -66.1‰ HITRAN: 38.4 ~~GEISA: -60.1‰~~ -102‰ -90.6‰ 8.98‰ -66.4‰ -91.3‰	GGG2014: -89.5‰ GGG2020: -73.0‰ HITRAN: 37.3‰ ~~GEISA: -67.6‰~~ -77.3‰ -66.1‰ 38.4 ‰ -60.1‰
Darwin	GGG2014: -126‰ GGG2020: -78.1‰ HITRAN: -60.1‰ GEISA: -72.1‰ ~~SEOM: -88.3‰~~ GGG2014 GGG2020 HITRAN GEISA SEOM	GGG2014: -107‰ GGG2020: -59.4‰ HITRAN: -39.9‰ ~~GEISA:-76.5‰~~ -126‰ -78.1‰ -60.1‰ -72.1‰ -88.3‰	GGG2014: -113‰ GGG2020: -60.8‰ HITRAN: -37.4‰ ~~GEISA: -82.8‰~~ -107‰ -59.4‰ -39.9‰ -76.5‰
Tsukuba	GGG2014: -125‰ GGG2020: -101‰ HITRAN: -52.7‰ GEISA: -71.1‰ ~~SEOM: -77.3‰~~ GGG2014 GGG2020 HITRAN GEISA SEOM	GGG2014: -105‰ GGG2020: -81.8‰ HITRAN: -31.3‰ ~~GEISA: -73.4‰~~ -125‰ -101‰ -52.7‰ -71.1‰ -77.3‰	GGG2014: -109‰ GGG2020: -81.8‰ HITRAN: -26.3‰ ~~GEISA: -76.7‰~~ -105‰ -81.8‰ -31.3‰ -73.4‰
Ascension Island	GGG2014: -130‰ GGG2020: -89.3‰ HITRAN: -94.2‰ GEISA: -70.6‰ ~~SEOM: -104‰~~ GGG2014 GGG2020 HITRAN GEISA SEOM	GGG2014: -113‰ GGG2020: -72.0‰ HITRAN: -76.1‰ ~~GEISA: -77.1‰~~ -130‰ -89.3‰ -94.2‰ -70.6‰ -104‰	GGG2014: -117‰ GGG2020: -72.6‰ HITRAN: -73.9‰ ~~GEISA: -83.8‰~~ -113‰ -72.0‰ -76.1‰ -77.1‰

Table 5. Averaged values of δ^{13}C from all TCCON sites for all possible combinations of $^{12}CH_4$ with window 4 of $^{13}CH_4$ for each spectral database.

Site	Database	Windows 1 & 4	Windows 2 & 4	W
Ny-Ålesund	GGG2014: 2.46‰ GGG2020: 9.08‰ HITRAN: -49.1‰ ~~GEISA: -26.6‰~~ GGG2014 GGG2020 HITRAN GEISA	GGG2014: 29.7‰ GGG2020: 36.3‰ HITRAN: -21.3‰ ~~GEISA: -20.0‰~~ 2.46‰ 9.08‰ -49.1‰ -26.6‰	GGG2014: 16.1‰ GGG2020: 28.6‰ HITRAN: -22.4‰ ~~GEISA: -27.9‰~~ 29.7‰ 36.3‰ -21.3‰ -20.0‰	GGG201 GGG202 HITRAN ~~GEISA~~
Darwin	GGG2014: -117‰ GGG2020: -131‰ HITRAN: 65.3‰ ~~GEISA: 134‰~~ GGG2014 GGG2020 HITRAN GEISA	GGG2014: -97.4‰ GGG2020: -114‰ HITRAN: 88.1‰ ~~GEISA: 129‰~~ -117‰ -131‰ 65.3‰ 134‰	GGG2014: -103‰ GGG2020: -115‰ HITRAN: 91.0‰ ~~GEISA: 121‰~~ -97.4‰ -114‰ 88.1‰ 129‰	GGG20 GGG20 HITRA ~~GEIS~~
Tsukuba	GGG2014: -21.8‰ GGG2020: -26.7‰ HITRAN: 14.5‰ ~~GEISA: 56.4‰~~ GGG2014 GGG2020 HITRAN GEISA	GGG2014: 0.953‰ GGG2020: -5.62‰ HITRAN: 37.5‰ ~~GEISA: 53.8‰~~ -21.8‰ -26.7‰ 14.5‰ 56.4‰	GGG2014: -4.12‰ GGG2020: -5.66‰ HITRAN: 42.9‰ ~~GEISA: 50.1‰~~ 0.953‰ -5.62‰ 37.5‰ 53.8‰	GGG201 GGG202 HITRA ~~GEISA~~
Ascension Island	GGG2014: -125‰ GGG2020: -140‰ HITRAN: 11.5‰ ~~GEISA: 87.4‰~~ GGG2014 GGG2020 HITRAN GEISA	GGG2014: -107‰ GGG2020: -124‰ HITRAN: 31.8‰ ~~GEISA: 79.9‰~~ -125‰ -140‰ 11.5‰ 87.4‰	GGG2014: -112‰ GGG2020: -124‰ HITRAN: 34.3‰ ~~GEISA: 72.0‰~~ -107‰ -124‰ 31.8‰ 79.9‰	GGG20 GGG20 HITRA ~~GEISA~~

~~Table ??~~ Tables 4 and 5 indicates a wide range of results, ~~which is unsurprising given that we calculate~~ suggesting either significant differences in spectroscopic parameters or large retrieval uncertainty. GGG identifies the mean uncertainty ~~on a~~ of $^{13}CH_4$ ~~retrieval over all datasets at between 18-25%~~retrievals to between 0.5 - 2 ppb (∼2.5-10%) depending on the database and TCCON site. However, given that ~~the values indicated in Table ?? are averages, this uncertainty~~ these uncertainties can be averaged over a long period of time, they should reduce significantly (by ∼ x200 in the case of Darwin), meaning that the precision ~~should be comparable to that~~ of ~~an individual~~ ~~12~~13CH_4 ~~retrieval. Firstly~~ retrievals should be very high (e.g. <0.006 ppb). Therefore precision errors cannot explain the differences in δ^{13}C values shown in Tables 4 and 5, meaning differences in the spectroscopic databases are the key sources of errors in $^{13}CH_4$ retrievals. This therefore suggests that knowledge of $^{13}CH_4$ retrievals spectroscopic parameters must be improved before serious attempts at remote sensing of $^{13}CH_4$ can be made.

Looking at these results in more detail, considering the retrievals that use $^{13}CH_4$ from window 1, these combinations yield surprisingly consistent results site to site and window to window~~, except~~. Except for the HITRAN results which show significant bias at the Ny-Ålesund site, although this is contrasted by the HITRAN results at Darwin and Tsukuba which indicates values very close to what might be expected. Indeed, the results from HITRAN at Ny-Ålesund ~~show results~~ are significantly different from those at any of the other sites, which can be explained by $^{13}CH_4$ retrievals showing significantly larger biases (at least x2 any other database), and a very high Pearson's correlation (0.7). The results from GEISA are the most consistent across the window combinations for all sites, showing a maximum of ∼13‰ variation across all window ~~combination~~ combinations which is a remarkable result. For comparison purposes, HITRAN shows ∼25‰ variation, GGG2020 ∼18‰ variation and GGG2014 ∼25-60‰ variation. The variations between databases in the same window combinations are larger than those in-between windows, suggesting variable dependence on local conditions, and thus differences in spectroscopic parameters. Except for HITRAN, all of the databases and windows seem to underestimate the accepted δ^{13}C values.

The results for the $^{13}CH_4$ window 4 combination are highly varied, more so than those shown for the window 1 combinations. For Ny-Ålesund both the GGG2014 and 2020 results show high levels of bias and significant variation, while the HITRAN and GEISA results show much lower bias levels and generally consistent results, with the HITRAN window 1 & 4 combination showing a realistic result. This is contrasted by the results from Darwin, where large biases are observed from all of the databases, but similar levels of consistency between the window combinations. Tsukuba again shows large bias levels between databases and windows, with GEISA showing high levels of consistency but large bias. Ascension Island shows similar results to Darwin (except for the HITRAN calculations), indicating similar sensitivity to background conditions. Pearson correlation coefficients for window 4 generally ~~indicate~~ indicated lower levels of sensitivity to variations of local conditions than window 1, suggesting the spectroscopic parameters for $^{13}CH_4$ in window 4 have significant uncertainty. ~~Further to this~~Furthermore, we found that the retrieval errors generated from $^{13}CH_4$ in window 4 were at least double those from $^{13}CH_4$ in window 1. This lower uncertainty is key in explaining the lower variation in δ^{13}C metric calculated using window 1.

Overall the results in ~~Table ??~~ Tables 4 and 5 suggest using GEISA2020 $^{13}CH_4$ retrievals from window 1 to calculate δ^{13}C values, showing high levels of consistency across all windows and sites, and relatively low bias levels. This consistency is surprising and is worth further investigation, however, Window 4 for all spectroscopic databases yield far less accurate results, suggesting more work must be done for spectroscopic parameters in this window for $^{13}CH_4$.

4 Discussion

We have shown the presence of correlations between variations in specific local conditions and retrieval biases in this paper, however it should be noted that other local conditions do vary in parallel with those indicated in Sect, 3.3. It is therefore likely that each window and spectroscopic database show bias variability due to the variation of a number of conditions simultaneously, which is why each TCCON site shows different results. The key message remains true however, that different windows in different spectroscopic databases are sensitive to varying degrees to local changing conditions. Further analysis in this topic should be assessed, for example the impact of the air-mass factor changes or variations in the O_2 retrievals may be important. We note Cygan et al. (2012); Ngo et al. (2013) identify Voigt broadening parameters for O_2 as insufficient. The release of the GGG2020 environment may allow for the testing of the impact of non-Voigt parameters on O_2 retrievals. Currently, the bias present in TCCON O_2 retrievals are removed by air-mass correction factors, based on results from the O_2 parameters in the GGG2014 database, and modified for each TCCON site. This means the use of DMFs for the comparisons in this study are likely to unfairly favour the results from the GGG2014 database. Therefore a potential option for comparison purposes would be to calculate volume mixing ratios (VMRs) based on dividing the retrieved methane quantity by the column of dry air calculated using the surface pressure and water vapour column, as opposed to the O_2 column. While this method would remove the biases associated with the GGG2014 O_2, it would introduce biases associated with the measured surface pressure, and the water vapour column which are more significant than the biases associated with O_2 (Wunch et al., 2011). Further, one of the key reasons for using DMFs as opposed to VMRs is that O_2 is well known and a constant, and can be used as a standard between all of the sites. Therefore, while the use of DMFs introduces biases, the use of VMRs would make the different sites less comparable.

We have also not considered errors in the instruments themselves, for example variations in the instrument line shape function between different TCCON instruments could cause additional biases.

We note that advancements are currently being tested on retrievals of methane from TCCON spectra, for example with the "SFIT4" algorithm (Zhou et al., 2019), which allows for profile retrievals and would therefore be less subject to the methane profile errors that can occur in GGG retrievals (Wunch et al., 2011). In addition to profile retrievals, this study used the GGG2014 retrieval software, while the more recent version of this software GGG2020 has also recently been ~~released~~announced. This update includes an improved spectroscopic database (this database was used in this study, wrapped in the GGG2014 software)~~and the ability to use~~, which includes non-Voigt line shapes for methane~~.~~, and possibly other gases. However, the GGG2014 software used in this study cannot leverage the non-Voigt parameters currently embedded in the GGG2020 spectroscopic database. Therefore further analysis using the GGG2020 software instead of GGG2014, and the use of other algorithms in this study could yield improved or different results. However, it is likely that the bias problems identified in this study may remain to some degree.

In addition to understanding the biases associated with retrieving $^{12}CH_4$ DMFs from TCCON spectra with differing spectroscopic databases, this study touches on a question that is of some interest to the community, ~~“can we~~ namely whether it is possible to calculate realistic and constant $\delta^{13}C$ values from TCCON~~”~~. The results shown in ~~Table ??~~ Tables 4 and 5 suggest not this is not yet possible, since they are often significantly different from the tropospheric average $\delta^{13}C$ value which is assumed to be -47‰ (Sherwood et al., 2016), and variable between databases and windows. There are some interesting cases where results close to the expected $\delta^{13}C$ value are calculated (e.g. windows 1 & 1 for HITRAN at Tsukuba), however given the same database in the same windows yields a completely inaccurate result at another TCCON site, it is challenging to draw any conclusions without further analysis. What is clear however, is that the $\delta^{13}C$ values ~~calculate~~ calculated using $^{13}CH_4$ retrievals from window 1 tend to have less biases than those calculated using window 4, and show less variation between windows and TCCON sites, as well as more consistent results between the spectroscopic databases. The implication of these results are that window 1 is superior to window 4 for retrieving $^{13}CH_4$ DMF, however whether this is due to superior information content, or more accurate knowledge of spectroscopic parameters requires further research.

However, given that TCCON retrieves total column estimates, and not in-situ samples as assumed by Sherwood et al. (2016), this assumption of -47‰ is a little unfair, since this is ~~an assumption~~ based on lower tropospheric averages, and does not take into account sink processes that occur further up into the atmosphere. For example Rigby et al. (2017) assume a -2.6‰ fractionation due to the chlorine sink in the stratosphere, and significant fractionation does occur in the troposphere with the

OH sink (Röckmann et al., 2011). However, it can be argued here that the priority in calculating an accurate value of δ^{13}C from TCCON is a full assessment of all of the systematic biases present in the retrievals, most notably the spectroscopic biases, before discussion of the true δ^{13}C value of the total column.

The results in this study have also shown the impact of water vapour is significant when considering inter-window and spectroscopic database retrievals, as identified by Figs. 11 and 9. Therefore further work is necessary to characterise the impacts on the biases exhibited in the results shown in this paper.

5 Conclusions

In this study, using the GGG2014 retrieval environment we retrieve $^{12}CH_4$ DMFs from four TCCON sites over the course of a year in each case, with the aim of understanding the biases associated with retrieving methane in the TROPOMI spectral region as opposed to standard TCCON methane windows. Four different windows covering the spectral range of the future S5/UVNS instrument and the current S5P/TROPOMI instrument are used. Three of the windows are routinely used in TCCON products, but the TROPOMI/UVNS window in the 4190-4340 cm^{-1} range is not. We use five sources of spectroscopic parameters, the HITRAN2016, GEISA2020, SEOM-IAS and internal TCCON databases (GGG2014 and GGG2020) in order to assess the impact of spectroscopic database uncertainties.

Firstly we analysied the quality of fit of each of the windows for each of the spectroscopic databases, for each window we find the GGG2020 spectroscopic database shows the best fit metrics, except in window 1, where the SEOM-IAS database has the best quality of fit. We note that while each TCCON site shows different fit statistics for each window, the order of the spectroscopic databases in terms of quality of fit remains the same in all cases, with GGG2020 showing the best, followed by GGG2014, HITRAN2016 and GEISA2020. Window 1 shows the poorest quality fit of all of the windows, indicating room to improve the spectroscopic parameters for window 1.

Using metrics based on bias w.r.t the standard TCCON methane retrieval window (a weighted average of three windows), we found that each of the TCCON sites, the GGG2014 and GGG2020 databases exhibited normalised biases <1 in the standard TCCON windows, meaning that these biases were below the retrieval noise limit and were therefore not significant. However in window 1, both GGG2014 and GGG2020 indicated biases >1 for most of the TCCON sites, suggesting TCCON retrievals in window 1 have a significant bias to the standard TCCON window. Similarly the HITRAN and GEISA databases showed significant biases (in some cases >2) w.r.t the standard in windows 1 and 4, indicating significant disagreement between the standard TCCON retrievals and the HITRAN and GEISA databases in these windows. Only the GEISA database showed significant disagreement with the standard in windows 2 & 3, which, based on the other results shown in this paper, suggest the GEISA database as having the largest differences of all of the databases considered in this study.

The sensitivity of the retrieved $^{12}CH_4$ DMFs to locally changing conditions such as water vapour, SZA and temperature is investigated. We find significant levels of dependence on these variations that are not necessarily mirrored across all of the TCCON sites. We conclude that some retrieval windows and spectroscopic databases are more sensitive to variable conditions than others. This sensitivity is exacerbated at TCCON locations with highly variable and challenging local conditions.

The δ^{13}C metric calculated in this study show significant bias w.r.t the expected total column value of -47‰. However, the use of the 4265 cm^{-1} window shows significant benefit over the 6076 cm^{-1} window, and more consistent results across spectroscopic databases. Yet, the high levels of differences between the spectroscopic databases suggest high levels of uncertainty in $^{13}CH_4$ parameters, and further work must be done to reduce these uncertainties.

The analysis in this study led to two key conclusions, firstly we recommend including the TROPOMI SWIR spectral region (in this study, window 1) into future TCCON methane retrievals. This is based on comparable fit statistics with the original TCCON methane windows, and the significant bias w.r.t the standard TCCON retrieval product. Secondly, the different spectral windows used to generate the TCCON methane products are affected by local condition variability to varying degrees. Suggesting the weighted average normally used to generate TCCON methane products should be a unique formation depending on TCCON site and season.

Code and data availability. The GGG2014 retrieval environment is available at https://tccon-wiki.caltech.edu, and TCCON L1b spectra are available upon discussion with the relevant site PI

Appendix A: Spectral Fits

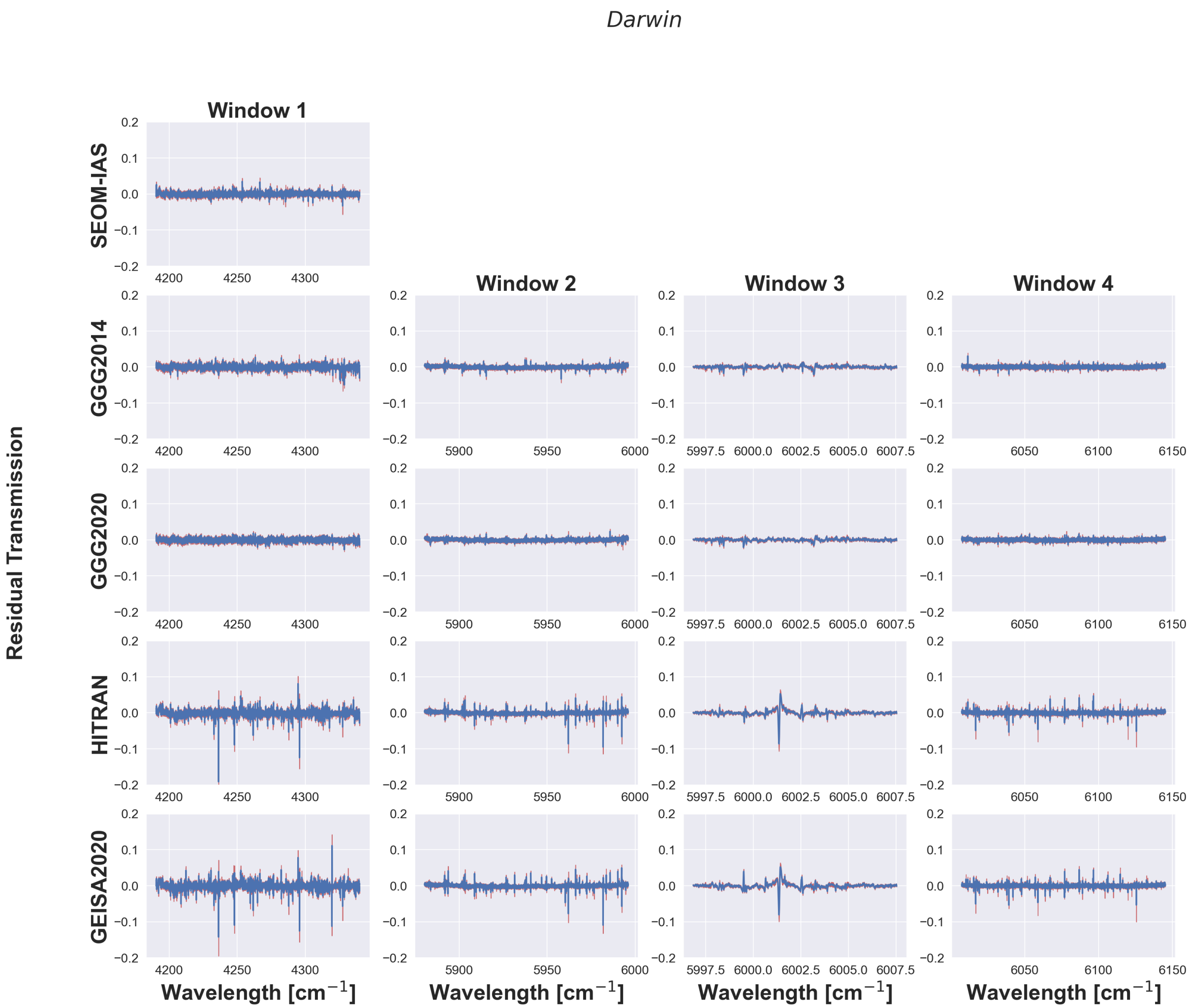


Figure A1. Example residual transmission spectra calculated from measured and fitted spectra from Darwin site in 2020.

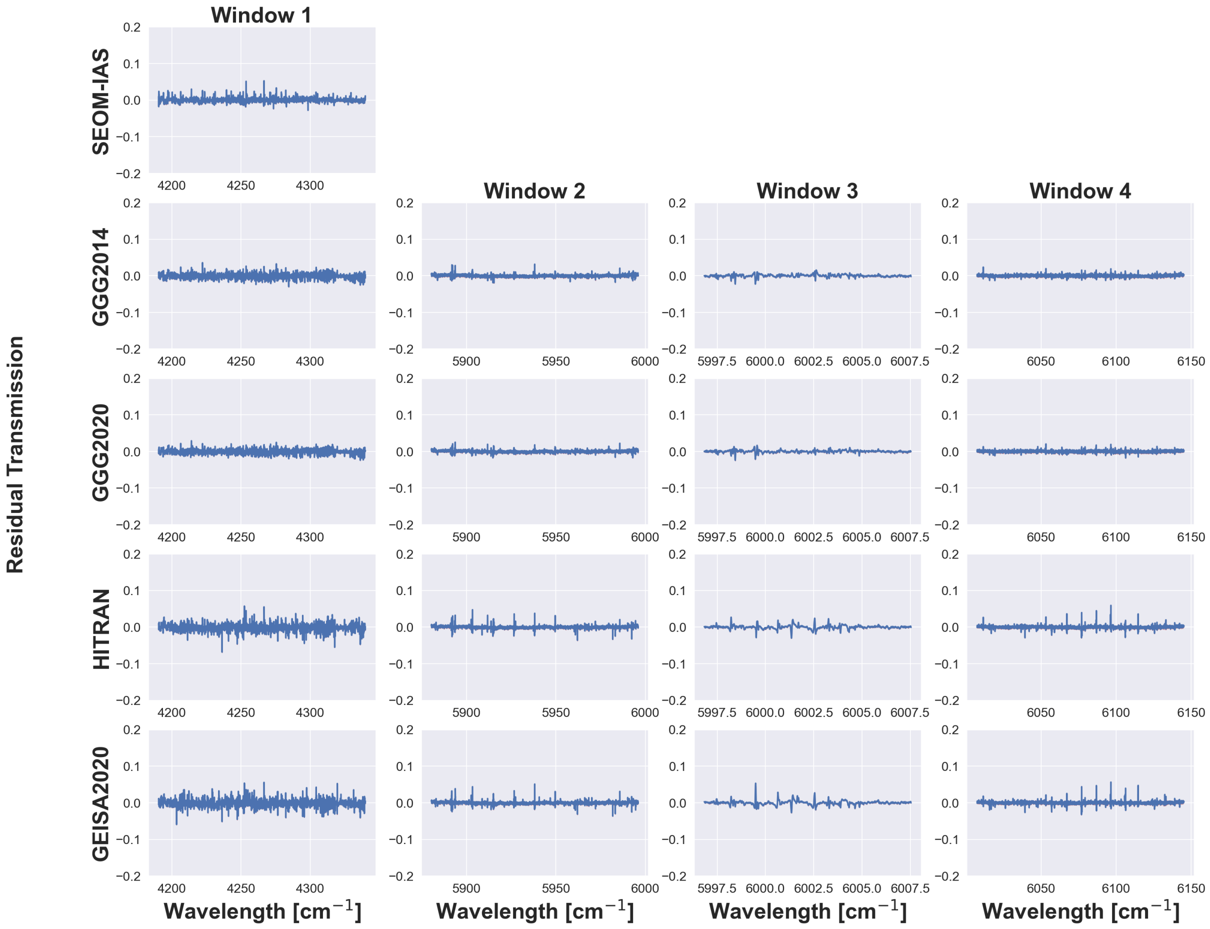


Figure A2. Example residual transmission spectra calculated from measured and fitted spectra from Tsukuba site in 2019.

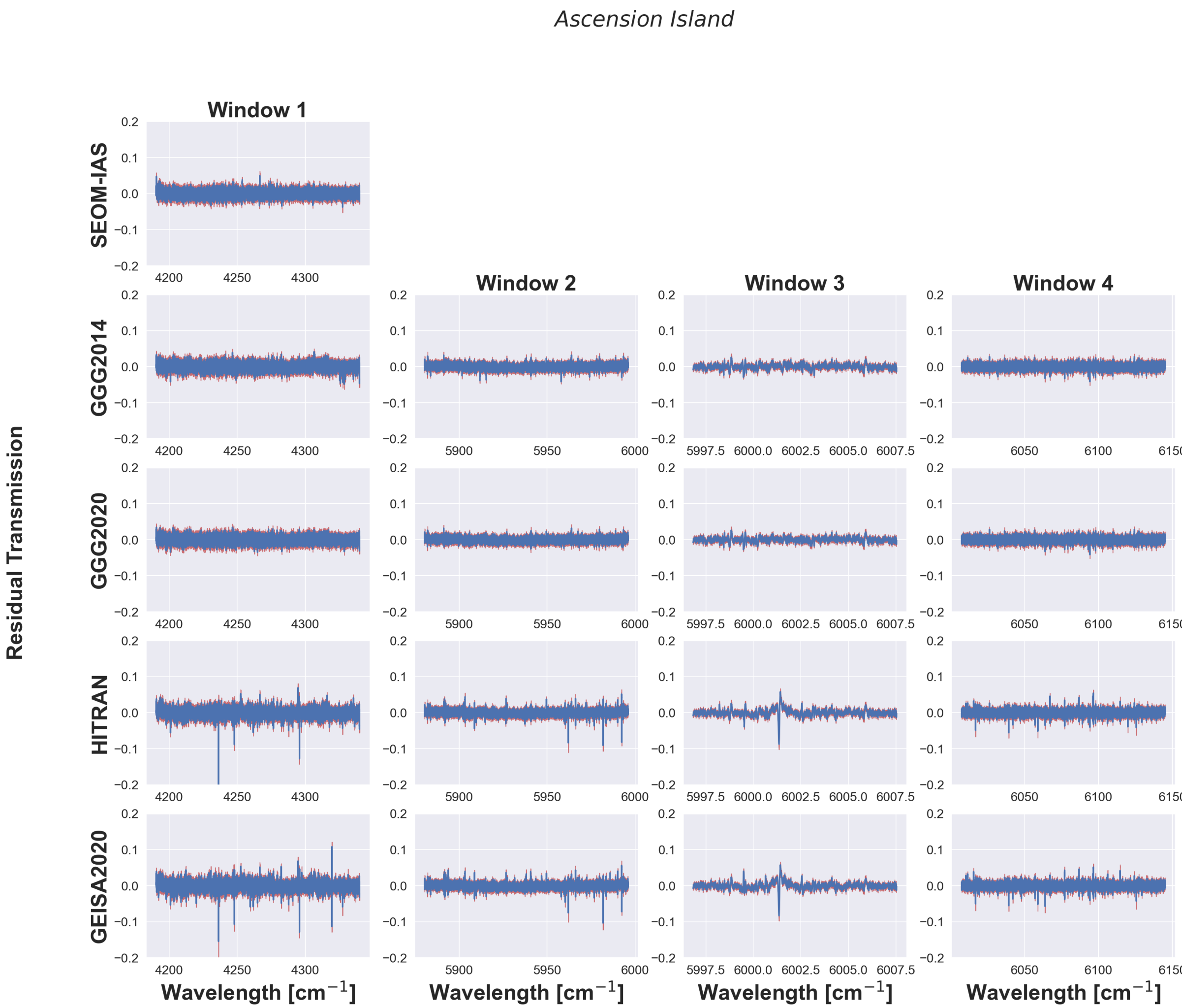


Figure A3. Example residual transmission spectra calculated from measured and fitted spectra from Ascension Island site in 2015.

Author contributions. MB provided and processed the Ny-Ålesund data, ND provided and processed the Darwin data, DGF provided and processed the Ascension Island TCCON data and IM provided and processed the Tsukuba TCCON data. EM devised and performed the study, analysed the data and wrote the paper. BV consulted on the interpretation of the results. All authors reviewed the paper.

Competing interests. BV is an associate editor for the joint (AMT/ACP) special issue "TROPOMI on Sentinel-5 Precursor: first year in operation". DGF is an associate editor for AMT.

Acknowledgements. This study has been performed in the framework of the postdoctoral Research Fellowship program of the European Space Agency (ESA). The GGG2014 retrieval environment was developed at the California institute of Technology (Caltech), and is available at https://tccon-wiki.caltech.edu. Thanks to Geoff Toon at Caltech for providing advice on the use of GGG2014. HITRAN2016 is available at https://hitran.org/, GEISA2020 is available from http://ara.abct.lmd.polytechnique.fr/index.php?page=geisa-2. The SEOM-IAS database is available at https://www.wdc.dlr.de/seom-ias/. The TCCON line lists are described in the GGG documentation. The Ascension Island TCCON station has been supported by ESA under grant 3-14737 and by the German Bundesministerium für Wirtschaft und Energie (BMWi) under grants 50EE1711C and 50EE1711E. We thankfully acknowledge funding from the University of Bremen and the German Research Foundation (DFG) – Projektnummer 268020496 – TRR 172 – (AC)3 and thank the AWIPEV station personnel in Ny-Ålesund and the Alfred-Wegener Institute Potsdam for their continuous support. The TCCON sites at Tsukuba is supported in part by the GOSAT series project. The Darwin TCCON station is funded by NASA grants NAG5- 12247 and NNG05-GD07G and supported by the Australian Research Council (ARC) grants DP1601001598, DP140101552, DP110103118, DP0879468 and LP0562346. Thanks to Anu Dudhia at Oxford University for the Fortran routine to convert the GEISA line structure into a HITRAN line structure.

References

(TCCON Wiki login) RunningGGG2020 < Main < TCCON Wiki, https://tccon-wiki.caltech.edu/Main/RunningGGG2020#Run_post_45Processing, 2022.

Albert, S., Bauerecker, S., Boudon, V., Brown, L., Champion, J.-P., Loëte, M., Nikitin, A., and Quack, M.: Global analysis of the high resolution infrared spectrum of methane 12CH4 in the region from 0 to 4800cm-1, Chemical Physics, 356, 131–146, https://doi.org/10.1016/j.chemphys.2008.10.019, 2009.

An, X., Caswell, A. W., and Sanders, S. T.: Quantifying the temperature sensitivity of practical spectra using a new spectroscopic quantity: Frequency-dependent lower-state energy, Journal of Quantitative Spectroscopy and Radiative Transfer, 112, 779–785, https://doi.org/10.1016/j.jqsrt.2010.10.014, 2011.

Armante, R., Scott, N., Crevoisier, C., Capelle, V., Crepeau, L., Jacquinet, N., and Chédin, A.: Evaluation of spectroscopic databases through radiative transfer simulations compared to observations. Application to the validation of GEISA 2015 with IASI and TCCON, Journal of Molecular Spectroscopy, 327, 180–192, https://doi.org/10.1016/j.jms.2016.04.004, 2016.

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P., Colin, R., DeCola, P., DeMazière, 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., 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, Geophysical Research Letters, 32, L15S01, https://doi.org/10.1029/2005GL022386, 2005.

Birk, M., Wagner, G., Loos, J., Mondelain, D., and Campargue, A.: ESA SEOM-IAS – Spectroscopic parameters database 2.3 μm region, https://doi.org/10.5281/ZENODO.1009126, 2017.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: Mission Objectives and Measurement Modes, Journal of the Atmospheric Sciences, 56, 127–150, https://doi.org/10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2, 1999.

Brown, L. R., Sung, K., Benner, D. C., Devi, V. M., Boudon, V., Gabard, T., Wenger, C., Campargue, A., Leshchishina, O., Kassi, S., Mondelain, D., Wang, L., Daumont, L., Régalia, L., Rey, M., Thomas, X., Tyuterev, V. G., Lyulin, O. M., Nikitin, A. V., Niederer, H. M., Albert, S., Bauerecker, S., Quack, M., O'Brien, J. J., Gordon, I. E., Rothman, L. S., Sasada, H., Coustenis, A., Smith, M. A., Carrington, T., Wang, X. G., Mantz, A. W., and Spickler, P. T.: Methane line parameters in the HITRAN2012 database, Journal of Quantitative Spectroscopy and Radiative Transfer, 130, 201–219, https://doi.org/10.1016/j.jqsrt.2013.06.020, 2013.

Buzan, E. M., Beale, C. A., Boone, C. D., and Bernath, P. F.: Global stratospheric measurements of the isotopologues of methane from the Atmospheric Chemistry Experiment Fourier transform spectrometer, Atmos. Meas. Tech, 9, 1095–1111, https://doi.org/10.5194/amt-9-1095-2016, 2016.

Checa-Garcia, R., Landgraf, J., Galli, A., Hase, F., Velazco, V. A., Tran, H., Boudon, V., Alkemade, F., and Butz, A.: Mapping spectroscopic uncertainties into prospective methane retrieval errors from Sentinel-5 and its precursor, Atmospheric Measurement Techniques, 8, 3617–3629, https://doi.org/10.5194/amt-8-3617-2015, 2015.

Crisp, D., Fisher, B. M., O 'dell, C., Frankenberg, C., Basilio, R., Bösch, H., Brown, L. R., Castano, R., Connor, B., Deutscher, N. M., Eldering, A., Griffith, D., Gunson, M., Kuze, A., Mandrake, L., Mcduffie, J., Messerschmidt, J., Miller, C. E., Morino, I., Natraj, V., Notholt, J., O 'brien, D. M., Oyafuso, F., Polonsky, I., Robinson, J., Salawitch, R., Sherlock, V., Smyth, M., Suto, H., Taylor, T. E., Thompson, D. R., Wennberg, P. O., Wunch, D., and Yung, Y. L.: The ACOS CO 2 retrieval algorithm – Part II: Global X CO 2 data characterization, Atmos. Meas. Tech, 5, 687–707, https://doi.org/10.5194/amt-5-687-2012, 2012.

Cygan, A., Lisak, D., Wójtewicz, S., Domysławska, J., Hodges, J. T., Trawiński, R. S., and Ciuryło, R.: High-signal-to-noise-ratio laser technique for accurate measurements of spectral line parameters, Physical Review A - Atomic, Molecular, and Optical Physics, 85, 022 508, https://doi.org/10.1103/PhysRevA.85.022508, 2012.

Delahaye, T., Armante, R., Scott, N., Jacquinet-Husson, N., Chédin, A., Crépeau, L., Crevoisier, C., Douet, V., Perrin, A., Barbe, A., Boudon, V., Campargue, A., Coudert, L., Ebert, V., Flaud, J.-M., Gamache, R., Jacquemart, D., Jolly, A., Kwabia Tchana, F., Kyuberis, A., Li, G., Lyulin, O., Manceron, L., Mikhailenko, S., Moazzen-Ahmadi, N., Müller, H., Naumenko, O., Nikitin, A., Perevalov, V., Richard, C., Starikova, E., Tashkun, S., Tyuterev, V., Vander Auwera, J., Vispoel, B., Yachmenev, A., and Yurchenko, S.: The 2020 edition of the GEISA spectroscopic database, Journal of Molecular Spectroscopy, 380, 111 510, https://doi.org/10.1016/j.jms.2021.111510, 2021.

Drummond, J. R. and Mand, G. S.: The measurements of pollution in the troposphere (MOPITT) instrument: Overall performance and calibration requirements, Journal of Atmospheric and Oceanic Technology, 13, 314–320, https://doi.org/10.1175/1520-0426(1996)013<0314:TMOPIT>2.0.CO;2, 1996.

Fisher, R. E., France, J. L., Lowry, D., Lanoisellé, M., Brownlow, R., Pyle, J. A., Cain, M., Warwick, N., Skiba, U. M., Drewer, J., Dinsmore, K. J., Leeson, S. R., Bauguitte, S. J.-B., Wellpott, A., O'Shea, S. J., Allen, G., Gallagher, M. W., Pitt, J., Percival, C. J., Bower, K., George, C., Hayman, G. D., Aalto, T., Lohila, A., Aurela, M., Laurila, T., Crill, P. M., McCalley, C. K., and Nisbet, E. G.: Measurement of the 13C isotopic signature of methane emissions from northern European wetlands, Global Biogeochemical Cycles, 31, 605–623, https://doi.org/10.1002/2016GB005504, 2017.

Galli, A., Butz, A., Scheepmaker, R. A., Hasekamp, O., Landgraf, J., Tol, P., Wunch, D., Deutscher, N. M., Toon, G. C., Wennberg, P. O., Griffith, D. W., and AbenI.: CH4, CO, and H2O spectroscopy for the Sentinel-5 Precursor mission: An assessment with the Total Carbon Column Observing Network measurements, Atmospheric Measurement Techniques, 5, 1387–1398, https://doi.org/10.5194/amt-5-1387-2012, 2012.

Gordon, I., Rothman, L., Hill, C., Kochanov, R., Tan, Y., Bernath, P., Birk, M., Boudon, V., Campargue, A., Chance, K., Drouin, B., Flaud, J.-M., Gamache, R., Hodges, J., Jacquemart, D., Perevalov, V., Perrin, A., Shine, K., Smith, M.-A., Tennyson, J., Toon, G., Tran, H., Tyuterev, V., Barbe, A., Császár, A., Devi, V., Furtenbacher, T., Harrison, J., Hartmann, J.-M., Jolly, A., Johnson, T., Karman, T., Kleiner, I., Kyuberis, A., Loos, J., Lyulin, O., Massie, S., Mikhailenko, S., Moazzen-Ahmadi, N., Müller, H., Naumenko, O., Nikitin, A., Polyansky, O., Rey, M., Rotger, M., Sharpe, S., Sung, K., Starikova, E., Tashkun, S., Auwera, J. V., Wagner, G., Wilzewski, J., Wcisło, P., Yu, S., and Zak, E.: The HITRAN2016 Molecular Spectroscopic Database, Journal of Quantitative Spectroscopy and Radiative Transfer, https://doi.org/10.1016/j.jqsrt.2017.06.038, 2017.

Hu, H., Hasekamp, O., Butz, A., Galli, A., Landgraf, J., Aan De Brugh, J., Borsdorff, T., Scheepmaker, R., and Aben, I.: The operational methane retrieval algorithm for TROPOMI, Atmos. Meas. Tech, 9, 5423–5440, https://doi.org/10.5194/amt-9-5423-2016, 2016.

Ingmann, P., Veihelmann, B., Langen, J., Lamarre, D., Stark, H., and Courrèges-Lacoste, G. B.: Requirements for the GMES Atmosphere Service and ESA's implementation concept: Sentinels-4/-5 and -5p, Remote Sensing of Environment, 120, 58–69, https://doi.org/10.1016/J.RSE.2012.01.023, 2012.

IPCC: Fifth Assessment Report - Impacts, Adaptation and Vulnerability, http://www.ipcc.ch/report/ar5/wg2/, 2014.

Jacquinet-Husson, N., Armante, R., Scott, N. A., Chédin, A., Crépeau, L., Boutammine, C., Bouhdaoui, A., Crevoisier, C., Capelle, V., Boonne, C., Poulet-Crovisier, N., Barbe, A., Chris Benner, D., Boudon, V., Brown, L. R., Buldyreva, J., Campargue, A., Coudert, L. H., Devi, V. M., Down, M. J., Drouin, B. J., Fayt, A., Fittschen, C., Flaud, J. M., Gamache, R. R., Harrison, J. J., Hill, C., Hodnebrog, Hu,

S. M., Jacquemart, D., Jolly, A., Jiménez, E., Lavrentieva, N. N., Liu, A. W., Lodi, L., Lyulin, O. M., Massie, S. T., Mikhailenko, S., Müller, H. S., Naumenko, O. V., Nikitin, A., Nielsen, C. J., Orphal, J., Perevalov, V. I., Perrin, A., Polovtseva, E., Predoi-Cross, A., Rotger, M., Ruth, A. A., Yu, S. S., Sung, K., Tashkun, S. A., Tennyson, J., Tyuterev, V. G., Vander Auwera, J., Voronin, B. A., and Makie, A.: The 2015 edition of the GEISA spectroscopic database, Journal of Molecular Spectroscopy, 327, 31–72, https://doi.org/10.1016/j.jms.2016.06.007, 2016.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., Bergamaschi, P., Bergmann, D., Blake, D. R., Bruhwiler, L., Cameron-Smith, P., Castaldi, S., Chevallier, F., Feng, L., Fraser, A., Heimann, M., Hodson, E. L., Houweling, S., Josse, B., Fraser, P. J., Krummel, P. B., Lamarque, J. F., Langenfelds, R. L., Le Quéré, C., Naik, V., O'doherty, S., Palmer, P. I., Pison, I., Plummer, D., Poulter, B., Prinn, R. G., Rigby, M., Ringeval, B., Santini, M., Schmidt, M., Shindell, D. T., Simpson, I. J., Spahni, R., Steele, L. P., Strode, S. A., Sudo, K., Szopa, S., Van Der Werf, G. R., Voulgarakis, A., Van Weele, M., Weiss, R. F., Williams, J. E., and Zeng, G.: Three decades of global methane sources and sinks, https://doi.org/10.1038/ngeo1955, 2013.

Laughner, J., Andrews, A., Roche, S., Kiel, M., and Toon, G.: ginput v1.0.10: GGG2020 prior profile software (Version 1.0.10), https://doi.org/10.22002/D1.1944, 2021.

Lorente, A., Borsdorff, T., Butz, A., Hasekamp, O., Aan De Brugh, J., Schneider, A., Wu, L., Hase, F., Kivi, R., Wunch, D., Pollard, D. F., Shiomi, K., Deutscher, N. M., Velazco, V. A., Roehl, C. M., Wennberg, P. O., Warneke, T., and Landgraf, J.: Methane retrieved from TROPOMI: Improvement of the data product and validation of the first 2 years of measurements, Atmospheric Measurement Techniques, 14, 665–684, https://doi.org/10.5194/amt-14-665-2021, 2021.

Malina, E., Yoshida, Y., Matsunaga, T., and Muller, J. P.: Information content analysis: The potential for methane isotopologue retrieval from GOSAT-2, Atmospheric Measurement Techniques, 11, 1159–1179, https://doi.org/10.5194/amt-11-1159-2018, 2018.

Malina, E., Hu, H., Landgraf, J., and Veihelmann, B.: A study of synthetic 13 CH 4 retrievals from TROPOMI and Sentinel-5/UVNS, Atmos. Meas. Tech, 12, 6273–6301, https://doi.org/10.5194/amt-12-6273-2019, 2019.

Mendonca, J., Strong, K., Sung, K., Devi, V. M., Toon, G. C., Wunch, D., and Franklin, J. E.: Using high-resolution laboratory and ground-based solar spectra to assess CH4 absorption coefficient calculations, Journal of Quantitative Spectroscopy and Radiative Transfer, 190, 48–59, https://doi.org/10.1016/j.jqsrt.2016.12.013, 2017.

Ngo, N. H., Lisak, D., Tran, H., and Hartmann, J. M.: An isolated line-shape model to go beyond the Voigt profile in spectroscopic databases and radiative transfer codes, Journal of Quantitative Spectroscopy and Radiative Transfer, 129, 89–100, https://doi.org/10.1016/j.jqsrt.2013.05.034, 2013.

Nikitin, A., Lyulin, O., Mikhailenko, S., Perevalov, V., Filippov, N., Grigoriev, I., Morino, I., Yoshida, Y., and Matsunaga, T.: GOSAT-2014 methane spectral line list, Journal of Quantitative Spectroscopy and Radiative Transfer, 154, 63–71, https://doi.org/10.1016/J.JQSRT.2014.12.003, 2015.

Nikitin, A., Chizhmakova, I., Rey, M., Tashkun, S., Kassi, S., Mondelain, D., Campargue, A., and Tyuterev, V.: Analysis of the absorption spectrum of 12CH4 in the region 5855–6250 cm-1 of the $2\nu3$ band, Journal of Quantitative Spectroscopy and Radiative Transfer, 203, 341–348, https://doi.org/10.1016/J.JQSRT.2017.05.014, 2017.

Nisbet, E. G., Dlugokencky, E. J., Manning, M. R., Lowry, D., Fisher, R. E., France, J. L., Michel, S. E., Miller, J. B., White, J. W. C., Vaughn, B., Bousquet, P., Pyle, J. A., Warwick, N. J., Cain, M., Brownlow, R., Zazzeri, G., Lanoisellé, M., Manning, A. C., Gloor, E., Worthy, D. E. J., Brunke, E.-G., Labuschagne, C., Wolff, E. W., and Ganesan, A. L.: Rising atmospheric methane: 2007-2014 growth and isotopic shift, Global Biogeochemical Cycles, 30, 1356–1370, https://doi.org/10.1002/2016GB005406, 2016.

Rella, C. W., Hoffnagle, J., He, Y., and Tajima, S.: Local-and regional-scale measurements of CH 4 , δ 13 CH 4 , and C 2 H 6 in the Uintah Basin using a mobile stable isotope analyzer, Atmos. Meas. Tech, 8, 4539–4559, https://doi.org/10.5194/amt-8-4539-2015, 2015.

Rigby, M., Montzka, S. A., Prinn, R. G., White, J. W. C., Young, D., O'Doherty, S., Lunt, M. F., Ganesan, A. L., Manning, A. J., Simmonds, P. G., Salameh, P. K., Harth, C. M., Mühle, J., Weiss, R. F., Fraser, P. J., Steele, L. P., Krummel, P. B., McCulloch, A., and Park, S.: Role of atmospheric oxidation in recent methane growth., Proceedings of the National Academy of Sciences of the United States of America, 114, 5373–5377, https://doi.org/10.1073/pnas.1616426114, 2017.

Röckmann, T., Brass, M., Borchers, R., and Engel, A.: The isotopic composition of methane in the stratosphere: High-altitude balloon sample measurements, Atmospheric Chemistry and Physics, 11, 13 287–13 304, https://doi.org/10.5194/acp-11-13287-2011, 2011.

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., Raymond, P. A., Dlugokencky, E. J., Houweling, S., Patra, P. K., Ciais, P., Arora, V. K., Bastviken, D., Bergamaschi, P., Blake, D. R., Brailsford, G., Bruhwiler, L., Carlson, K. M., Carrol, M., Castaldi, S., Chandra, N., Crevoisier, C., Crill, P. M., Covey, K., Curry, C. L., Etiope, G., Frankenberg, C., Gedney, N., Hegglin, M. I., Höglund-Isakson, L., Hugelius, G., Ishizawa, M., Ito, A., Janssens-Maenhout, G., Jensen, K. M., Joos, F., Kleinen, T., Krummel, P. B., Langenfelds, R. L., Laruelle, G. G., Liu, L., Machida, T., Maksyutov, S., McDonald, K. C., McNorton, J., Miller, P. A., Melton, J. R., Morino, I., Müller, J., Murgia-Flores, F., Naik, V., Niwa, Y., Noce, S., O'Doherty, S., Parker, R. J., Peng, C., Peng, S., Peters, G. P., Prigent, C., Prinn, R., Ramonet, M., Regnier, P., Riley, W. J., Rosentreter, J. A., Segers, A., Simpson, I. J., Shi, H., Smith, S. J., Steele, P. L., Thornton, B. F., Tian, H., Tohjima, Y., Tubiello, F. N., Tsuruta, A., Viovy, N., Voulgarakis, A., Weber, T. S., van Weele, M., van der Werf, G. R., Weiss, R. F., Worthy, D., Wunch, D., Yin, Y., Yoshida, Y., Zhang, W., Zhang, Z., Zhao, Y., Zheng, B., Zhu, Q., Zhu, Q., and Zhuang, Q.: The Global Methane Budget 2000-2017, Earth System Science Data Discussions, pp. 1–138, https://doi.org/10.5194/essd-2019-128, 2019.

Scheepmaker, R. A., Aan De Brugh, J., Hu, H., Borsdorff, T., Frankenberg, C., Risi, C., Hasekamp, O., Aben, I., and Landgraf, J.: HDO and H2O total column retrievals from TROPOMI shortwave infrared measurements, Atmospheric Measurement Techniques, 9, 3921–3937, https://doi.org/10.5194/amt-9-3921-2016, 2016.

Schneising, O., Buchwitz, M., Reuter, M., Bovensmann, H., Burrows, J. P., Borsdorff, T., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Hermans, C., Iraci, L. T., Kivi, R., Landgraf, J., Morino, I., Notholt, J., Petri, C., Pollard, D. F., Roche, S., Shiomi, K., Strong, K., Sussmann, R., Velazco, V. A., Warneke, T., and Wunch, D.: A scientific algorithm to simultaneously retrieve carbon monoxide and methane from TROPOMI onboard Sentinel-5 Precursor, Atmospheric Measurement Techniques Discussions, pp. 1–44, https://doi.org/10.5194/amt-2019-243, 2019.

Sherwood, O., Schwietzke, S., Arling, V., and Etiope, G.: Global Inventory of Fossil and Non-fossil Methane δ13C Source Signature Measurements for Improved Atmospheric Modeling, https://doi.org/10.15138/G37P4D, 2016.

TCCON: Using the cc option in GGG2014 - Tccon-wiki, https://tccon-wiki.caltech.edu/index.php?title=Software/GGG/Download/GGG{_}2014{_}Release{_}Notes/Using{_}the{_}cc{_}option{_}in{_}GGG2014{&}highlight=continuum+curvature, 2020.

Toon, G. C.: Atmospheric Line List for the 2014 TCCON Data Release, https://doi.org/10.14291/tccon.ggg2014.atm.R0/1221656, 2015.

Tran, H., Ngo, N., and Hartmann, J.-M.: Efficient computation of some speed-dependent isolated line profiles, Journal of Quantitative Spectroscopy and Radiative Transfer, 129, 199–203, https://doi.org/10.1016/J.JQSRT.2013.06.015, 2013.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027, 2012.

Weidmann, D., Hoffmann, A., Macleod, N., Middleton, K., Kurtz, J., Barraclough, S., and Griffin, D.: The Methane Isotopologues by Solar Occultation (MISO) Nanosatellite Mission: Spectral Channel Optimization and Early Performance Analysis, Remote Sensing, 9, 1073, https://doi.org/10.3390/rs9101073, 2017.

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., and Hurst, D. F.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmos. Meas. Tech, 3, 1351–1362, https://doi.org/10.5194/amt-3-1351-2010, 2010.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, Phil. Trans. R. Soc. A, 369, 2087–2112, https://doi.org/10.1098/rsta.2010.0240, 2011.

Wunch, D., Toon, G., Sherlock, V., Deutscher, N., Liu, C., Feist, D., and Wennberg, P.: The Total Carbon Column Observing Network's GGG2014 Data Version, Tech. rep., https://doi.org/10.14291/tccon.ggg2014.documentation.R0/1221662, 2015.

Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I., and Yokota, T.: Retrieval algorithm for CO 2 and CH 4 column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite, Atmos. Meas. Tech, 4, 717–734, https://doi.org/10.5194/amt-4-717-2011, 2011.

Yoshida, Y., Kikuchi, N., Morino, I., Uchino, O., Oshchepkov, S., Bril, A., Saeki, T., Schutgens, N., Toon, G. C., Wunch, D., Roehl, C. M., Wennberg, P. O., Griffith, D. W. T., Deutscher, N. M., Warneke, T., Notholt, J., Robinson, J., Sherlock, V., Connor, B., Rettinger, M., Sussmann, R., Ahonen, P., Heikkinen, P., Kyrö, E., Mendonca, J., Strong, K., Hase, F., Dohe, S., and Yokota, T.: Improvement of the retrieval algorithm for GOSAT SWIR XCO2 and XCH4 and their validation using TCCON data, Atmospheric Measurement Techniques, 6, 1533–1547, https://doi.org/10.5194/amt-6-1533-2013, 2013.

Zhou, M., Langerock, B., Sha, M. K., Kumps, N., Hermans, C., Petri, C., Warneke, T., Chen, H., Metzger, J. M., Kivi, R., Heikkinen, P., Ramonet, M., and De Mazière, M.: Retrieval of atmospheric CH4 vertical information from ground-based FTS near-infrared spectra, Atmospheric Measurement Techniques, 12, 6125–6141, https://doi.org/10.5194/amt-12-6125-2019, 2019.