

This is a single pdf that includes point-by-point responses to the reviews and a marked-up manuscript version showing the changes to the original manuscript.

Below we have included the full text of the review (black text), interspersed with our responses addressing the reviewer's specific comments and changes to the manuscript in italicised blue font.

Reviewer 1 Comments:

G1 – The word “calibrate” has been used throughout to describe reducing the mismatch between TCCON and EM27/SUN retrievals, when in most cases “scale” would be a better choice. Calibration is generally reserved for something directly observed of known accuracy, and XCO₂ is neither. In some situations “calibration” is okay so the wording is not too awkward (e.g., calibration curve), but should be changed in most instances.

We agree and have made the changes, for example in Section 4.1.2. we made changes to a few sentences:

“From the EM27 and TCCON measurements in 2016, we derived a ~~calibration~~ scaling factor for the EM27. We ~~calibrated~~ scaled the EM27 measurements to Wollongong TCCON data because the Wollongong XCO₂ has been calibrated against aircraft profile measurements that are traceable to the WMO in-situ scale (Wunch et al., 2010).”

*“From this exercise, we arrive at a ~~calibration~~ scaling factor of $EM27=0.9954*TCCON$.”*

“Therefore, we are confident that the ~~calibration~~ scaling factor relative to TCCON was consistent during the campaign with that derived relative to Wollongong TCCON.”

Also under Section 4.3

*“From here, we derived a ~~calibration~~ scaling factor of $0.9927*EM27$ for GOSAT H-gain and $0.9983*EM27$ for the GOSAT M-gain,...”*

G2 – It seems that differences in averaging kernels (AKs) and a priori profiles have not been considered in this study. E.g., on P2L35 it is stated measurements can be compared directly. It should be explicitly stated why AKs were not considered. Perhaps the a priori profiles are close enough to the true profiles *in this location* that accounting for these different sensitivities would make little difference? Also, it seems observations were first averaged rather than directly compared from both the EM27/SUN and GOSAT? It would also be interesting to know how much of the H to M gain bias is from AKs, if any.

We've added to Section 4.1.2, 2nd paragraph:

“We did not consider the averaging kernels in this work. The averaging kernels of the EM27 have been previously presented and compared to TCCON in a study by Hedelius et al. (2016). In their study, they found that although there are differences in the TCCON and EM27 averaging kernels, the effect of the differences in averaging kernels from the top of the atmosphere cancel out the effect of differences at the bottom.”

We have also reworded the last sentence in Section 4th paragraph:

“The Alice Springs measurements are unique considering that they have been collected from a clean desert environment and can be directly compared to GOSAT retrievals using the M-gain measurements.”

To

The Alice Springs measurements are unique considering that they have been collected from a clean desert environment where GOSAT M-gain soundings are also abundant and close enough (<100 km) to compare with the EM27.

Specific comments

S1, P1L7: State length of campaign here

Done, thanks

S2, P1L_1: How exactly are the different gains chosen and used for GOSAT? Is gain chosen in real time by an onboard sensor, or are all gains recorded and the retrieval just picks the best one later?

We have revised the last sentences of the first paragraph in the Introduction:

“These gain settings are prespecified at certain locations because GOSAT does not observe both M and H gains simultaneously. For the majority of the soundings over land, H-gain is used. GOSAT M-gain retrievals over land are used over surfaces that are bright in the SWIR such as deserts and semi-arid regions. However, a bias between GOSAT M- and H-gain retrievals of XCO₂ has been reported, along with a lack of M-gain validation with TCCON (Yoshida et al., 2013). In Australia where surface reflectivity values are generally high, GOSAT was configured to observe using M gain for much of the land surface, at first to avoid detector saturation. However, it was found that some observations using H gain did not result in saturated signals and were still useful. Therefore from 2012-Feb-11, GOSAT started to perform alternate observations using both H and M gains in order to investigate the differences from retrievals between these two gain settings. Therefore from 2012-Feb-11, GOSAT started to perform alternate observations using both H and M gains in order to investigate the differences from retrievals between these two gain settings.”

How/why did the gains change in Fig. 1 going from v 2.60 to v 2.72?

This is a good question, thanks. So under Section 2 “Alice Springs Australia Site Description”, paragraph 1, we’ve added:

“Note that for Ver. 02.72 FTS SWIR L2 retrievals, upgraded input and reference products were used. For example, there was an improvement in the spatial resolution of the cloud flagging procedure, which employs the CAI L2 (Cloud and Aerosol Imager, Level 2) data. This improvement resulted in better screening of the data and may have resulted in an increased number of soundings that were passed for the Ver. 02.72 FTS SWIR L2 retrievals.”

These changes are described in the release notes:

https://data2.gosat.nies.go.jp/doc/documents/ReleaseNote_FTSSWIRL2_V02.72_en.pdf

S3, P2L16: What does this mean? There are still some large cities in deserts.

True, this sentence was poorly constructed, we now changed it to:

“The world’s deserts and semi-arid regions encompass large areas that are mostly undisturbed by anthropogenic emissions and are important for understanding the carbon cycle.”

S4, P2L20: This goes back to the question of how gain is chosen, but are M-gain regions always exclusively semi-arid? If so where does the classification of climate come from in the algorithm?

Thanks for the question. We believe we’ve addressed this above

S5, P2L20: Carbon cycle studies could be (and have been) carried out with biased satellite retrievals, even if the bias is “small.” I think the focus here though is if the goal is to continuously improve the accuracy of these studies, then accuracy of satellite retrievals needs to be improved as well. What would happen if all observations over high albedo were not available? Likely results would be biased, and would lead to misinterpretation.

This is correct! Thank you for this very important point. We would like to use it in the manuscript so we added:

“Recently, the importance of semi-arid regions in the carbon cycle inter-annual variability has been highlighted (Poulter et al., 2014). If the goal is to continuously improve the accuracy of carbon cycle studies, then the accuracy of satellite retrievals needs to be improved as well because if observations over high albedo were not available or biased, flux estimates would likely be biased as well and would lead to misinterpretation. Therefore, measurements over M-gain regions are needed by the satellite community (Yoshida et al., 2013) and highlights the significance of this study.

S6, P2L25: Why is high reflectivity a challenge for satellite observations? How does the extra reflectance influence the retrievals negatively? (Same question on P15L12) S7, P2L25: How does having challenging observations naturally lead to their improvement? (Maybe the meaning here is it rather an opportunity?)

Thanks for these good questions we would like to address them at the same time. In the manuscript, we explained this in the text as below:

“The desert environment provides high surface reflectivity, which challenges satellite retrievals because aerosols, depending on type, can lead to an effect called optical path lengthening and this effect is dominant in regions with high albedo (Yoshida et al., 2013). Recently, Iwasaki et al. (2019) showed that there is an increase in the XCO₂ retrievals using their PPDF-S algorithm when the albedo at 1.6μm was high, implying that the retrieved XCO₂ is strongly related to the surface albedo. This challenge leads to the improvement of satellite retrievals.”

S8, P2L27: Why are such studies needed?

Thanks, we have expanded this sentence to:

“Benchmark measurements and pilot studies for desert sites are needed to assess the benefit and feasibility of such sites because setting up a TCCON site in remote deserts will be difficult logistically and financially.”

S9, P3L1-2: Suggest you pick a notation for EM27 or EM27/SUN early on and stay consistent throughout the entire paper.

Good point, we have done this early on in the Introduction by writing:

“We address this need by utilizing a well-established portable spectrometer, an EM27/SUN by Bruker Optics GmbH (Gisi et al., 2012; Frey et al., 2015; Hase et al., 2015; Hedelius et al., 2016; Frey et al., 2018), which was retrofitted with a protective fairing and automated solar tracker hatch for operations in a harsh environment. For brevity, we will refer to this instrument as EM27.”

S10, P3L2: Are v 2.60 data bias corrected? How?

Actually, both are NOT bias corrected, so we rewrote that part:

“We compare NIES GOSAT retrieval versions 2.60 and the new version 2.72, which are both not bias corrected.”

S11, P3L1-2: Suggest you move this sentence to later on in the paragraph where again it states EM27/SUN retrievals were compared with GOSAT.

Thanks, we moved this after mentioning Section 4.

S12, P3L9: How was the EM27/SUN retrofitted?

We added this in the introduction:

“...retrofitted with a protective fairing and automated solar tracker hatch for operations in a harsh environment.”

S13, P3L10: suggest you include 2 decimals on latitude. Also, include dates here.

Done, thanks.

S14, P4L4: Specify these are column measurements.

Well spotted, thanks.

S15, P4L7: Was the 2nd detector present for this study? Give full spectral range for this detector.

No, it wasn't present so we deleted that sentence, which takes care of the next issue

S16, P4L7-8: Suggest you move this sentence before the previous one. Currently it sounds like O2 is measured on the secondary detector.

True indeed, thanks, we did this.

S17, P4L9: What is the spectral resolution of TCCON measurements?

TCCON uses a maximum optical path difference (MOPD) of 45 cm, corresponding to a spectral resolution of 0.02 cm^{-1} . We've added it now, thanks.

S18, P5L2: "greenhouse gas (CO2 and CH4) total columns" could simply be replaced with "XCO2 and XCH4."

Thanks, done.

S19, P5L_23: Suggest you describe how Xair can provide info on stability (e.g., as a measure of retrieved O2, which is not particularly variable in dry air).

Thank you, this is a very good and important point. We improved the Xair description according to this suggestion.

"Xair is a good indicator of instrument stability and changes in spectrometer alignment because 1) VC_air is calculated using the surface pressure, which is independently measured by a pressure sensor to better than 0.3 hPa, keeping accuracy over long periods (Wunch et al., 2011); and 2) the atmospheric oxygen column is not particularly variable in dry air, hence the retrieved VC_O2 by the spectrometer should be close to constant. From Eq. 3, it follows that a perfectly accurate measurement would lead to an Xair value of unity, however, due to spectroscopic limitations, the actual value is approximately 0.98 for all TCCON sites.

S20, P6L2: Clarify "specific" here. Is it only soundings within 100 km of the EM27? Are they compared individually or averaged together?

We elaborated this paragraph to:

"The TANSO-FTS on GOSAT has a two-axis pointing system. Normally, TANSO-FTS follows an M shaped grid on a 5-point cross-track scan mode. By using this pointing system to vary the observation geometry, it is able to observe specific points, i.e. it can view targets with angles up to $\pm 20^\circ$ along the satellite track and by $\pm 30^\circ$ across the track. Specific point observations over Alice Springs were requested from July 2016, in preparation for the campaign in September. Five locations within 100 km of the center of Alice Springs were targeted (see Fig. 3).

S21, P7L4: Frey et al., 2018 (<https://doi.org/10.5194/amt-2018-146>) would also be an appropriate reference here.

Thanks, done.

S22, P8: Was there any other alignment of the EM27 or TCCON instrument during this period?

Good question. In the revised version, we state:

“We note that there were no re-alignments done on the EM27 and TCCON instruments during this period, however the EM27 clamshell cover and fairing were fitted on the EM27 in February 2016, it is possible that this may have affected the alignment resulting in a shift in the Xair.”

S23, P10 Fig6: Specify what points represent in caption. Daily averages? Daily averages within certain sza?

In the Figure caption, we now write:

“Weighted hourly averaged XCO₂ from TCCON vs XCO₂ from EM27 in Wollongong for 2015-2016. To avoid noisy data, only measurements corresponding to TCCON with retrieved Xair values within 0.9783 and 0.9853 were selected (Xair values within the 2nd to 98th percentile)”

S24, P11L2: How were GOSAT data interpolated? (spatially? temporally? method?)

We changed this sentence. The word “interpolated” was removed because it can be confusing. Instead we write:

“To demonstrate the variability of XCO₂ in the region, we plot in Fig. 8 the daily mean time series of GOSAT soundings using these two coincidence criteria: 1) all soundings within a 1000-km radius centered at the BOM facility in Alice Springs and 2) all soundings measured within the same day (local time). Each data point considered in the calculation of the daily mean is weighted by the corresponding reported retrieval error for that particular GOSAT sounding.”

S25, P11L3: This is an unweighted mean?

No, as mentioned above, each data point considered in the calculation of the daily mean is weighted by the corresponding retrieval error for each GOSAT sounding falling within the spatial and temporal coincidence criteria. This associated error is reported in the GOSAT data.

S26, P11L12: Quantify approximate magnitude of annual increase here.

We have quantified from the slope of line in Fig. 8 the following XCO₂ annual increase:

H-gain: 2.2815 ppm/year; intercept: 132.9996

M-Gain: 2.2829 ppm/year; intercept: 134.9174

So, at the end of the first paragraph in Sec. 4.3, we wrote:

“We confirm that the M-gain retrievals are biased high (around 2 ppm) compared to the H-gain retrievals. Nevertheless, both least squares fitted lines to the H and M-gain retrievals show an increase of about 2.28 ppm/year with y-intercept values at 133.0 ppm (H-gain) and 134.92 ppm (M-gain).”

S27, P12L1: Did the number of H-gain observations decrease, or just the number of successful retrievals? If it's just the retrievals could it be an increased failure of convergence?

The number of successful retrievals increased. At the end of Sec 2, we added:

“there was an improvement in the spatial resolution of the cloud flagging procedure, which employs the CAI L2 (Cloud and Aerosol Imager, Level 2) data. This improvement resulted in better screening of the data and may have resulted in an increased number of soundings that were passed for the Ver. 02.72 FTS SWIR L2 retrievals.”

S28, P12L2: Apr-Aug mean in absolute number of soundings would also be useful here, same on line 4.

We agree and included this (please refer to the reply below).

S29, P12L1-7: Seems a better sentence order could be: less rain -> less vegetation -> bright surface/more M-gain (rather than less vegetation -> bright surface/more M-gain -> less rain).

Thank you, we agree. This paragraph has now been changed to:

Rainfall records indicate that April-August 2011 had the least amount of rainfall on record in Alice Springs after 2002, which was only broken by 2015, then 2017 (http://www.bom.gov.au/climate/data/station_number_015590). This significant absence of rain in the region could result in stunted vegetation growth or mortality, which could have led to brighter surfaces. Bright surfaces mean more M-gain measurements and this may explain the 74% increase in M-gain soundings in April-August 2011 compared to the April-August 2010-2017 average of c. 220 soundings/month. In contrast, the number of H-gain retrievals seems to have diminished around 2011, coinciding with the dry months starting from April 2011. The number of H-gain soundings from April-August 2011 was 23% fewer compared to the April-August average from the years 2010-2017 (c. 450 soundings/month).

S30, P12Table1: What do the seconds mean on the measurement times? Start time? Central time? Seems the measurements could take up to 20 seconds. Also, the purpose of the first and second to last rows is not clear if no GOSAT soundings were acquired.

We removed the seconds and simply wrote:

The times correspond to the times when the satellite is directly above the site. The satellite normally performs 5 observations points across track, with an interferometric scan time lasting 4 seconds (Shiomi et al. 2006). But specific point observations deviate from this pattern by pointing and maximizing observations near the target.

S31, P13Fig9: I do not really like lines fit through single points. Granted the intercept is forced through zero, but I think this information would be better for a table.

Yes, the intercept is forced through zero because both instruments are expected to produce zero XCO₂ when there is no CO₂ absorption. We added the sentence to Sec. 4.1.2:

“The retrieval method is predicted to be both linear and have zero intercept (Wunch et al., 2010). Therefore, we fit hourly mean data from TCCON and EM27 employing linear least squares and force a zero intercept.”

We would like to keep the figure because for future campaigns and comparisons with more data points we would like to use this as reference.

S32, P14L2: Where did these values come from?

We clarified this item in the paper and also considered the standard errors of the weighted mean for version 2.72:

“The standard errors of the weighted hourly mean GOSAT specific point observations are: 0.25 ppm and 0.26 ppm for M-gains (versions 2.6 and 2.72, respectively) and 0.29 ppm and 0.28 ppm for H-Gains (versions 2.6 and 2.72, respectively). These values are taken from the averaged standard errors of the weighted hourly means from all specific point observation data falling within 100 km from the site for 2 Sep. 2015 - Feb. 2017”.

S33, P14L7: satellite measurements -> satellite and ground-based measurements

Done, thanks.

S34, P15L2: What about Wollongong? Darwin? Reunion? Showa? Arrival Heights? Ascension? (Fig. 1: <https://doi.org/10.5194/amt-9-2381-2016>)

We elaborated this sentence to:

“With the exception of Lauder, New Zealand, there are no permanently dedicated ground targets for satellite calibration and validation in the Southern Hemisphere that is far enough from water (to avoid a possible land-ocean bias in retrievals) and surrounded by homogeneous topography (to avoid a possible altitude bias).”

S35, P15L4-5: This sentence seems a bit redundant with the first sentence.

Agreed, we removed this sentence

S36, P16L_1: Also add a description on how GOSAT data can be acquired

Done. thanks

Technical

T1, P1L7: values, a -> values, another

done, thanks.

T2, P1L11: improve -> better understand (or estimated -> estimation of)

thanks, changed to “better understand”

T3, P2L4: along with -> and there has been

agreed, thanks.

T4, P2L6: version 2.72 -> version 2.72 Xgas retrieval algorithm (w/o sentence seems to be missing a subject)

Thanks, sentence modified.

T5, P2L8: precise -> precise and accurate (?)

Yes, that is right, change done.

T6, P2L17: by anthropogenic -> by recent anthropogenic

Thanks, this is better.

T7, P2L20: are -> is

Done.

T8, P2L25: provides -> has

Done, thanks.

T9, P2L25: which challenges -> which is a challenge for

Corrected, thank you.

T10, P2L27: could be -> is

Changed.

T11, P2L28: portable -> portable solar-viewing

Done, thanks.

T12, P3L8: “in the urban area” seems redundant, maybe omit?

Agreed, done.

T13, P3L17: average high -> average daily high

Done, thanks

T14, P3L18-19: omit parenthetical comment (already on page 2, and Fig 1 caption)
We agree and omission done.

T15, P3L21: maybe omit “reasonably accessible” as this is somewhat vague
We agree and did this.

T16, P5L19-20: suggest “s” on Ps and “a” on Na be subscripts
Changed, thanks.

T17, P5L21: can -> is (?)
Yes, thanks. The sentence has been changed.

T18, P6L6: retrievals -> retrievals separately.
This is better, thanks

T19, P8L6: should give -> gives
Done, thanks.

T20, P11L5: were -> are
Done, thanks.

T21, P12L1: H-gain -> the number of H-gain (or better, -> there were 23% fewer Hgain)

T22, P12L2: Could omit “derived”

Thanks, to accommodate both of these, we changed the sentences to:

“In contrast, the number of H-gain retrievals seems to have diminished around 2011, coinciding with the dry months starting from April 2011. H-gain soundings from April-August 2011 were 23% fewer compared to the April-August mean from the years 2010-2017.”

T23, P14L14: Could omit “in Alice Springs”
Thanks, done.

T24, P14L15: cover -> enclosure (x2)
Done, thanks.

T25, P14L17: omit “the assumption”
Omitted, thanks

T26, P15L8: suggest omission of “and on other future satellites”
Suggestion accepted, thanks.

T27, P15L11: provides -> has
Done.

T28, P15L12: to -> for
Done, thanks

T29, P15L12: this data is especially needed -> observations are especially needed here
This is an improvement, thanks

T30, P16L7: calculated -> characterized
Done, thanks.

T31, P16L9: co-funded -> co-acquired funding for (?)
Agreed, thanks

T32, P16L14: define RA

Defined, thanks

T33, P16L14: advise -> advice

Corrected, thanks.

Other notes/optional

O1, P2L15: The population of the greater LA area includes parts of other counties. <https://www.citypopulation.de/world/Agglomerations.html> lists the population as 17.7 million.

Thank you very much for this information. The agglomerated population count represents the region better, so we changed it according to this.

O2, P5L3: The authors may also consider publishing a description of the design and/or control software in the future (e.g., compare <https://doi.org/10.5194/amt-11-2173-2018>). Such a project/paper could be especially useful to the community if the control software were open source, and fully automated neither of which have been done yet. Example automation of OPUS: <https://doi.org/10.1364/AO.57.000689>, example of alternate solar tracking software: <http://hdl.handle.net/10222/64642>, Chapter 4.

Yes indeed, the EM27 enclosure has undergone several revisions, our aim is to have 3D drawings that we can share to the community. The Davis weather station has a built-in programmable logic; we are also planning to distribute the configuration.

O3, P6Fig2: More details on this schematic could be useful, such as the path of light, and parts that move/rotate or disassemble (it looks like a seam at the “v” part?) An actual picture at Alice Springs could be nice if available. A picture would provide the readers an idea if any precautions needed to be taken to prevent interference from unique Australian fauna, such as fencing or placement away from trees.

Yes, there is a seam at the “V” part for easy access to the solar beam inlet. However, the design has changed a little bit, so we are currently working on the engineering drawings that we can distribute. We have added a picture of the EM27 at the Alice Springs site as an inset to Fig 3. Fortunately, this site is just beside the airport and is fully fenced, so the noise and the fencing keep away large Australian fauna (e.g. Dingoes, Kangaroos, Sand goannas, etc.).

O4, P9L9-11: The ME at MOPD values seem particularly small, compared to typical values around 98-99% (Frey et al, 2018). It may be beneficial to realign the spectrometer. Though changes in Xair do not look large. . .

Yes, well noted. Unfortunately, unlike the 125HR, alignment of the EM27 is still not well documented in the community, therefore for a re-alignment we have to send the EM27 to Karlsruhe Institute of Technology (to Frey et al.) or directly to Bruker Germany. This was done in 2017.

O5, P10: Future measurements in Alice Springs may be useful to help derive an airmass correction for all EM27/SUN instruments. This dataset may not be ideal though since the ME at MOPD seems to differ from most other EM27/SUN instruments, and is only for part of one season.

Deriving an airmass correction is an excellent point. For future campaigns, we will make sure to have the EM27 re-aligned at Bruker or at Karlsruhe. A loan from the COCCON network (or others) may be a possibility as well.

O6, P14L2: A histogram of the standard deviations for the different gains could be useful.

This is a good idea. We followed this and added an inset to Fig 10. Also, we've added on Section 4.4, 2nd paragraph, sentences 1 to 3:

“Fig. 10 shows a plot of how the statistical errors in the estimated bias improve with the number of weeks in the campaign. The inset shows the normalized histogram of the reported GOSAT single sounding errors within 100 km of the site. Note that although H-gain retrievals have smaller errors on average, the amount of M-gain retrievals improves the standard error of the weighted mean.”

Reviewer 2 Specific comments

It appears that the Authors have not smoothed the XCO₂ retrievals from the various instruments to account for differences in the averaging kernels and priors (if these differ between the retrieval schemes) per the method of Rogers and Connor (2003) as implemented by Wunch et. al. (2010). It may be that the effect of this process would be negligible, however this should still be discussed and quantified.

Thank you for this important comment. Under Section 4.1, we've added:

"We did not consider the averaging kernels in this work. The averaging kernels of the EM27 have been previously presented and compared to TCCON in a study by Hedelius et al. (2016). In their study, they found that although there are differences in the TCCON and EM27 averaging kernels, the effect of the differences in averaging kernels from the top of the atmosphere cancel out the effect of differences at the bottom. Further work and more measurements may be necessary to better understand the cause of this phenomenon but this is beyond the scope of this study."

Introduction and Fig. 1. Is there a reason why there appear to be more v2.72 M-Gain retrievals?

Yes, this is a good question, thanks. So under Section 2 "Alice Springs Australia Site Description", paragraph 1, we've added:

"Note that for Ver. 02.72 FTS SWIR L2 retrievals, upgraded input and reference products were used. For example, there was an improvement in the spatial resolution of the cloud flagging procedure, which employs the CAI L2 (Cloud and Aerosol Imager, Level 2) data. This improvement resulted in better screening of the data and may have resulted in an increased number of soundings that were passed for the Ver. 02.72 FTS SWIR L2 retrievals."

These changes are described in the release notes:

https://data2.gosat.nies.go.jp/doc/documents/ReleaseNote_FTSSWIRL2_V02.72_en.pdf

There should be some commentary on how the GOSAT specific point observations differ from normal GOSAT observations in Section 3.3.

Good point, therefore under Section 3.3, we wrote:

"The TANSO-FTS on GOSAT has a two-axis pointing system. Normally, TANSO-FTS follows an M shaped grid on a 5-point cross-track scan mode. By using this pointing system to vary the observation geometry, it is able to observe specific points, i.e. it can view targets with angles up to $\pm 20^\circ$ along the satellite track and by $\pm 30^\circ$ across the track. Specific point observations over Alice Springs were requested from July 2016, in preparation for the campaign in September. Five locations within 100 km of the center of Alice Springs were targeted (see Fig. 3).

In section 4.1.1 comparing the retrievals of Xair between the EM27 and TCCON instruments, it is noted that there are both airmass and seasonal variations in Xair. It seems likely that at least some of the seasonal dependence might be caused by the differing ranges of solar zenith angles that are observed throughout the year. It would be interesting to plot Fig 5 for a limited range of solar zenith angles to identify if the airmass dependence is the only reason for the seasonal variation.

Thanks, this is a good point. We revised Fig. 5 to show Xair values at $SZA > 45^\circ$, $SZA < 45^\circ$ and all SZA and revised the section accordingly. The SZA dependence is small and within approximately 1%.

In any case, it would be useful to have a few more details about the year of inter-comparison measurements in the introductory paragraph of Sect. 4.1 e.g. number of days of measurements,

total number of measurements from each instrument and whether they were operated for the same periods on each day.

Under Section 3.2 we added in the last paragraph:

“In Wollongong, both the TCCON and EM27 solar tracker covers were opened and closed by the same pneumatic system. The same weather station provided the meteorological data that were used to pre-filter the data (e.g. fractional variation in solar intensity, wind speed and direction, pressure, etc). We did not filter the data according to solar zenith angles anymore, instead in addition to the pre-filter, we filtered out noisy retrievals by selecting only those with Xair values within 0.5 and 1.5 because anything beyond that would be unrealistic in the atmosphere but most likely be the cause of an interference or obstruction.”

And in Section 4.1, first paragraph, third sentence, we wrote:

“Here, we focus on comparisons of Xair and XCO₂ from measurements spanning almost one year (Nov. 2015 to Sep. 2016)” under varied environmental conditions. Both instruments normally measure at the same time, apart from interruptions due to occasional mid-infrared measurements with the 125HR or rare software glitches (e.g. JAVA issues).”

Section 4.1.2 and Fig. 6, what time averaging is applied to the data in Fig. 6? The colour scale and the size of the 1 σ uncertainties suggest daily averages, but this is not made explicit.

Actually, these are hourly averages, so in the text we wrote:

*“Therefore, we fit hourly mean data from TCCON and EM27 employing linear least squares and force a zero intercept. The standard errors of the weighted means are used as weights in the fit. From this exercise, we arrive at a scaling factor of EM27=0.9954*TCCON”*

There also appears to be quite a large variety in the magnitude of the 1 σ values in Fig. 6. which would bear explanation. Several times during the manuscript calibration factors are presented in the form INST1 = F * INST2 it would be useful to have some indication of the uncertainty of the F value, or the goodness of the regression fit used to derive it. Similarly, for the drift mentioned at P9L17.

We’ve added error bars on the inset plot for the drift and a Pearson R correlation coefficient (R=1) for the least squares fit in Fig. 6. The uncertainty in the F value arising from the linear fit, while forcing the intercept to zero, is very small (<1x10e-6), so we did not include it anymore.

Section 4.4. describing the length of campaign required to reduce the error in the bias estimates is interesting but would benefit from some discussion of what the target for accuracy and precision in the bias estimate should be.

Thanks, this is a good point. We wrote 0.1 ppm as reference and we cite the work of Miller at al., 2007 and added on Sec 4.4, 2nd paragraph:

“Miller (2007) showed that a comparison of surface CO₂ concentration data and XCO₂ data flux inversions clearly reveals a land-ocean bias in the XCO₂ retrievals, even when the bias is only 0.1 ppm”.

Conclusions: this section should have some comment relating to the presented dataset specifically. Also, as an interested reader it would be good to know if there are any plans to repeat the field campaign and build the dataset.

We included in the conclusion, 2nd to last sentence:

“Subject to funding and interest from the community, a repeat of the campaign is possible.”

P15L3, the sentence starting “With the exception of Lauder...” should be qualified.

Thanks, we reworded this whole sentence to:

“With the exception of Lauder, New Zealand, there are no permanently dedicated ground targets for satellite calibration and validation in the Southern Hemisphere that is far enough from water (to avoid a possible land-ocean bias in retrievals) and surrounded by homogeneous topography (to avoid a possible altitude bias)”

Data availability: what about the GOSAT data?

We included in the section “Data Availability”

GOSAT data can be obtained from the GOSAT Data Archive Service (GDAS) after registration for access via: https://data2.gosat.nies.go.jp/index_en.html

Technical corrections

P2L24 “atmospheric conditions that can serve as a calibration point...”

Yes, this is better, thanks.

P4 Fig. 1 caption: Start with “Location of..” to indicate that these do not represent the retrieved values, for consistency, use XCO₂ and XCH₄.

Good point, done. Thanks.

P4L7 “below 5000 cm⁻¹ which allows for...”

We improved the sentences related to this to:

“For total column measurements of CO₂, CH₄, H₂O and O₂ spectra in the near infrared, the instrument is fitted with an Indium Gallium Arsenide (InGaAs) detector dedicated to 5,000–12,000 cm⁻¹. This also enables measurements of spectra covering the O₂ bands necessary to derive column-averaged dry-air mole fractions of CO₂ and CH₄ similar to the method used by TCCON”

P5L20 The last sentence on this page is disjointed and difficult to read, consider revising.

Thanks, we agree and revised these sentences.

P6L4 “However, to construct the time-series...”

Done, thanks

P6L4 should refer to Fig. 8, not Fig. 7.

Well spotted, thanks.

P8 Fig. 4 caption: it should be sufficient to say month number and delete “in a year”.

Done, thanks

P9L12 Sentence can be finished after “campaign” to avoid repetition.

Agreed, we’ve done this. Thanks.

P10L7 “...retrievals results in lower...”

Done, thank you.

P11 Fig. 7 caption: use XCO₂ for consistency.

Done, thanks.

P11 Fig. 7 caption: "days on the x-axis are not"
Corrected, thank you.

P11 Fig. 7 caption: pluralise interruption
Corrected, thank you.

P12L2 Start the sentence with "The number of..." to clarify that the retrieved values are not reducing.
Thanks for the suggestion, we corrected it.

P16L14 Explain what RA means.
Done, thanks.

P18L4 correct the rendering of the subscript
We tried many times but there seems to be a difficulty in the subscripting between Overleaf and Mendeley. We will bring this up during the typesetting.

P18L33 The DOI is repeated
Corrected, thanks

P19L6 insert a space between inverse and models
Done, thanks.

Satellite and Ground-based Measurements of X_{CO_2} in a Remote Semi-Arid Region of Australia

Voltaire A. Velazco ¹, Nicholas M. Deutscher ¹, Isamu Morino ², Osamu Uchino ², Beata Bukosa ¹, Masataka Ajiro ², Akihide Kamei ², Nicholas B. Jones ¹, Clare Paton-Walsh ¹, and David W. T. Griffith ¹

¹Centre for Atmospheric Chemistry, School of Earth, Atmospheric and Life Sciences, Faculty of Science, Medicine and Health, University of Wollongong, Australia

²Satellite Remote Sensing Section and Satellite Observation Center, Center for Global Environmental Research, National Institute for Environmental Studies (NIES), Onogawa 16-2, Tsukuba, Ibaraki 305-8506, Japan

Correspondence: Voltaire A. Velazco (voltaire@uow.edu.au)

Abstract. In this study, we present ground-based measurements of column-averaged dry-air mole fractions (DMFs) of CO_2 (or X_{CO_2}) ~~from taken in a semi-arid region of Australia with~~ an EM27/SUN portable spectrometer, ~~equipped with an automated clam-shell cover, taken in a semi-arid region of Australia~~ clamshell cover. We compared these measurements to space-based X_{CO_2} retrievals from the Greenhouse Gases Observing Satellite (GOSAT). Side-by-side measurements of EM27/SUN with the Total Carbon Column Observing Network (TCCON) instrument at the University of Wollongong were conducted in 2015-2016 to derive an X_{CO_2} calibration scaling factor of 0.9954 relative to TCCON. Although we found a slight drift of ~~0.125% per month~~ 1.3% over three months in the calibration curve of the EM27/SUN vs TCCON X_{CO_2} , the alignment of the EM27/SUN proved stable enough for a two-week campaign, keeping the retrieved X_{air} values, ~~a~~ another measure of stability, to within 0.5% and the modulation efficiency to within 2%. From the measurements in Alice Springs, we confirm a small bias of around 2 ppm in the GOSAT M-gain to H-gain X_{CO_2} retrievals, as reported by the NIES GOSAT validation team. Based on the reported random errors from GOSAT, we estimate the required duration of a future campaign in order to ~~improve~~ better understand the estimated bias between the EM27/SUN and GOSAT. The dataset from the Alice Springs measurements is accessible at <http://dx.doi.org/10.4225/48/5b21f16ce69bc> (Velazco et al., 2018).

1 Introduction

The Greenhouse Gases Observing Satellite (GOSAT), launched on January 23, 2009, is the first satellite in orbit dedicated to the measurements of the two major anthropogenic greenhouse gases that cause global warming: CO_2 and CH_4 (Yokota et al., 2009; Kuze et al., 2009). Since its launch, GOSAT data have been used to study and determine the global distributions of CO_2 and CH_4 (Reuter et al., 2014; Houweling et al., 2015; Parker et al., 2015; Turner et al., 2015). The spectrometer on board GOSAT, called Thermal and Near infrared Sensor for Carbon ~~Observation-Fourier~~ Observation - Fourier Transform Spectrometer (TANSO-FTS), is able to simultaneously detect short wave infrared (SWIR: bands 1 at $0.76 \mu\text{m}$, 2 at $1.6 \mu\text{m}$ and 3 at $2.0 \mu\text{m}$) and thermal infrared (TIR: band 4 from 5.5 to $14.3 \mu\text{m}$). The TANSO-FTS has a footprint of about 10.5 km in diameter on the ground viewed at nadir, which is observed for 4 seconds. Bands 1, 2, and 3 are measured in two linear polar-

izations simultaneously with three selectable gains; high (H), medium (M) and low (L) (Suto et al., 2013). GOSAT employs the different gain settings to compensate for the different signal levels due to the reflective properties (albedo) of the Earth's surface, which also depend on wavelength. These gain settings are prespecified at certain locations because GOSAT does not observe both M and H gains simultaneously. For the majority of the soundings over land, H-gain is used. GOSAT M-gain retrievals over land are used over surfaces that are bright in the SWIR such as deserts and semi-arid regions. However, a bias between GOSAT M- and H-gain retrievals of X_{CO_2} has been reported, ~~along with and there has been~~ a lack of M-gain validation with TCCON ~~Yoshida et al. (2013).~~ (Yoshida et al., 2013). In Australia where surface reflectivity values are generally high, GOSAT was configured to observe using M gain for much of the land surface, at first to avoid detector saturation. However, it was found that some observations using H gain did not result in saturated signals and were still useful. Therefore from 2012-Feb-11, GOSAT started to perform alternate observations using both H and M gains in order to investigate the differences from retrievals between these two gain settings.

Fig. 1 shows the locations of GOSAT M-gain soundings over a period of one year obtained from the NIES (National Institute for Environmental Studies, Japan) version 2.72 X_{CO_2} retrieval algorithm (white squares) and data from the previous version 2.60 (black dots) for reference. Satellite-based retrievals of X_{CO_2} are validated by the Total Carbon Column Observing Network (TCCON), a ground-based network dedicated to the precise and accurate measurements of greenhouse gases (Wunch et al., 2011). TCCON data are used to validate satellite instruments such as GOSAT (Morino et al., 2011), GOSAT-2 (Matsunaga et al., 2018), OCO-2 ~~Crisp et al. (2017)~~ (Crisp et al., 2017) and Sentinel 5P (Borsdorff et al., 2018). However, apart from locations in the western U.S.A. (Dryden and Los Alamos TCCON sites), there are no operational TCCON stations (magenta stars in Fig. 1) that are ideally located in a clean desert or semi-arid region not influenced by nearby anthropogenic emissions and have plenty of nearby GOSAT M-gain soundings. The Dryden TCCON station at the Armstrong Flight Research Center (AFRC), Edwards, CA is located in the Mojave desert at 34.960° N, 117.881° W (700 m a.s.l). Although not as densely populated, AFRC is only approximately 100 km north Los Angeles (~~L.A. county~~ pop. c. ~~10.16-17.8~~ million) and 100 km east of Bakersfield, CA (pop. c. 376,380).

~~Although significantly less populated, the~~ The world's deserts and semi-arid regions encompass large areas that are mostly undisturbed by recent anthropogenic emissions and are important for understanding the carbon cycle. Recently, the importance of semi-arid regions in the carbon cycle inter-annual variability has been highlighted (Poulter et al., 2014; Ahlström et al., 2015; Trudinger et al., 2016). If the goal is to continuously improve the accuracy of carbon cycle studies, then the accuracy of satellite retrievals needs to be improved as well, because if observations over high albedo are not available or biased, flux estimates would likely be biased as well and would lead to misinterpretation. Therefore, measurements over M-gain regions are needed by the satellite community (Yoshida et al., 2013) ~~because understanding and minimizing biases in satellite retrievals over semi-arid regions are necessary for carbon cycle studies. This and~~ highlights the significance of this study.

Central Australia has a large semi-arid region that is relatively easy to access. There are plenty of uninhabited, vast, homogeneous areas that can accommodate large footprints of satellite-based sensors such as GOSAT, GOSAT-2, OCO-2, and Sentinel 5P. Central Australia is also an ideal place to measure undisturbed atmospheric conditions that can serve as calibration point for satellite retrievals of atmospheric composition. At the same time, the desert environment ~~provides~~ has

high surface reflectivity, which ~~challenges satellite retrievals and thus leads to their improvement~~ is a challenge for satellite retrievals because aerosols, depending on type, can lead to an effect called optical path lengthening and this effect is dominant in regions with high albedo (Yoshida et al., 2013). Recently, Iwasaki et al. (2019) showed that there is an increase in the X_{CO_2} retrievals using their PPDF-S algorithm when the albedo at $1.6 \mu\text{m}$ was high, implying that the retrieved X_{CO_2} is strongly related to the surface albedo. This challenge leads to the improvement of satellite retrievals. Despite the importance of desert locations like Central Australia in remote sensing, infrastructure support is not available and accessibility ~~could be~~ is a challenge. Benchmark measurements and pilot studies for desert sites are needed to assess the benefit and feasibility of such sites because setting up a TCCON site in remote deserts will be difficult logistically and financially. We address this need by utilizing a well-established portable solar-viewing spectrometer, an EM27/SUN by Bruker Optics GmbH (Gisi et al., 2012; Frey et al., 2015; Hase et al., 2015; Hedelius et al., 2016; ?), which we retrofitted (Gisi et al., 2012; Frey et al., 2015; Hase et al., 2015; Hedelius et al., 2016; ?), which was retrofitted with a protective fairing and automated solar tracker clamshell cover for operations in a harsh environment. For brevity, we will refer to this instrument as EM27. The instrument can measure spectra covering the spectral bands in the near infrared necessary to derive column-averaged dry-air mole fractions of CO_2 , CH_4 and CO with sufficient stability for short-term campaigns (Hedelius et al., 2016). The instrument was transported to Alice Springs in Central Australia with the primary objective of validating the GOSAT X_{CO_2} signal and making benchmark measurements in the region. The Alice Springs measurements are unique considering that they have been collected from a clean desert environment and can be directly compared to GOSAT retrievals using the where GOSAT M-gain measurements.

In this paper, we focus on measurements from the EM27 and GOSAT. We compare NIES GOSAT retrieval versions 2.60 and the new version 2.72, which are not bias corrected, with the retrievals from the soundings are also abundant and close enough (within 100 km) to compare with the EM27/SUN.

This manuscript is organized as follows. A description of Alice Springs and desert Australia is given in Section 2. In Section 3, we briefly discuss the instruments and methods, which are already well established. Results of measurement comparisons with the TCCON station in Wollongong and comparisons with the GOSAT M-gain and H-gain soundings in Alice Springs are shown and discussed under Section 4. We focus on X_{CO_2} measurements from the EM27 and GOSAT. We compare NIES GOSAT retrieval versions 2.60 and the new version 2.72, which are both not bias corrected, with the retrievals from the EM27. Statistical calculations and recommendations for a future Alice Springs campaign are also under Section 4. We provide our conclusions in Section 5.

2 Alice Springs Australia Site Description

We conducted a measurement campaign with the retrofitted EM27 /SUN system at the Australian Bureau of Meteorology (BOM) facility in Alice Springs ($23.723.79^\circ\text{S}$, $133.87133.89^\circ\text{E}$) near the end of September to the beginning of October from September 29 to October 6, 2016, with the primary objective of validating the GOSAT X_{CO_2} signal in this semi-arid region. Alice Springs is located in central Central Australia, also called the "Red Centre". Next to Darwin and Palmerston, it is the third-largest town in the Northern Territory of Australia with a population of 23,726 in the urban area (2016 census). The

vegetation around Alice Springs is composed mostly of dry scrubby grassland. The Alice Springs terrain, consisting mostly of sandy plains with some areas of rocky highland, is bounded by several deserts; the Tanami desert to the north, Simpson desert to the east and southeast, the Great Victoria desert to the south and the Gibson desert to the west. In September, statistical data from the Australian Bureau of Meteorology (BOM) show an average monthly rainfall of 8.7 mm, average [daily](#) high and low temperatures of 27.3 °C and 10.3 °C respectively and mean monthly sunshine of 300 hours. Collected GOSAT M-gain soundings over land for one whole year (~~black dots for version 2.6 and white squares for version 2.72~~) shown in Fig. 1 include a large part of this "Red Centre" including Alice Springs. [Note that for Ver. 02.72 FTS SWIR L2 retrievals, upgraded input and reference products were used. For example, there was an improvement in the spatial resolution of the cloud flagging procedure, which employs the CAI L2 \(Cloud and Aerosol Imager, Level 2\) data. This improvement resulted in better screening of the data and may have resulted in an increased number of soundings that were passed for the Ver. 02.72 FTS SWIR L2 retrievals.](#) Apart from remote regions in North Africa, the Middle East and near-densely-populated areas in California USA, Central Australia is the only ~~reasonably-accessible~~ region that provides ~~a rich~~ [an abundant](#) amount of M-gain soundings.

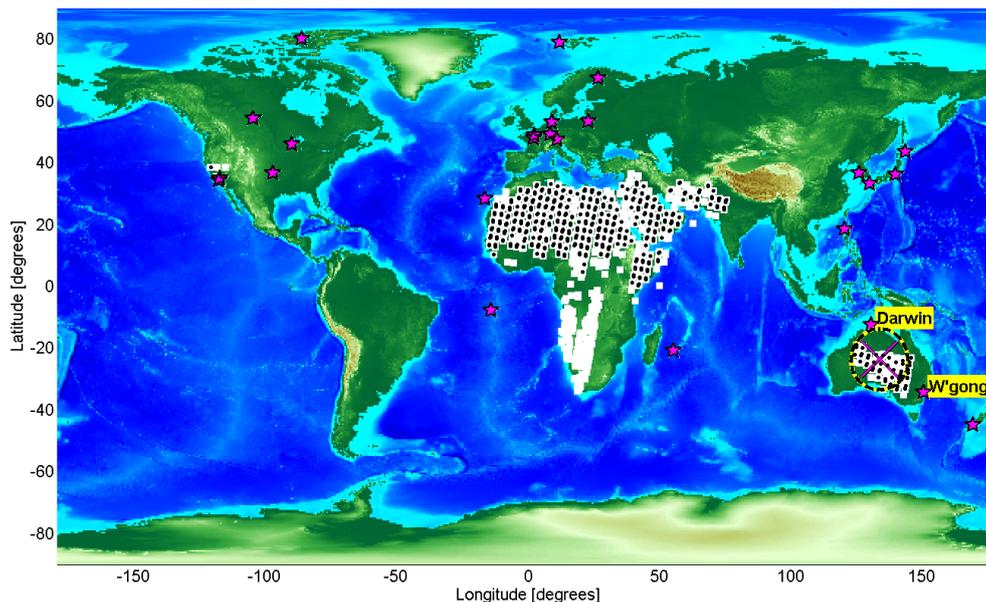


Figure 1. [Locations of GOSAT M-gain retrievals of soundings](#) for 2016. Black dots are from version 2.60 and the white squares are from version 2.72. Magenta stars indicate the location of operational TCCON sites as of October 2018. The ~~center of the~~ magenta cross marks the location of Alice Springs in Australia and the magenta circle has a radius of approximately 1000 km centered on Alice Springs. The Darwin and Wollongong (W'gong) sites in Australia are labeled.

3 Instruments and Methods

3.1 EM27/SUN with Automated Clamshell Cover

The EM27 /SUN system, which we will refer to as "EM27" from hereon system and its characterization are thoroughly described in the works of Gisi et al. (2012); Frey et al. (2015); Hase et al. (2015) and Hedelius et al. (2016). For total column measurements of CO₂, CH₄, H₂O and O₂ spectra in the near infrared, the instrument is fitted with an Indium Gallium Arsenide (InGaAs) detector dedicated to 5,000–12,000 cm⁻¹. In addition, a second extended InGaAs detector has subsequently been added for frequencies below 5000 cm⁻¹ allows for measurements of and N₂O. This also enables measurements of spectra covering the O₂-O₂ bands necessary to derive column-averaged dry-air mole fractions of CO₂ and CH₄ similar to the method used by TCCON. TCCON uses a maximum optical path difference (MOPD) of 45 cm, corresponding to a spectral resolution of 0.02 cm⁻¹. However, in contrast to TCCON, we used the typical maximum optical path difference (MOPD) EM27 we used has the typical MOPD of 1.8 cm, corresponding to a spectral resolution of 0.5 cm⁻¹.

We equipped the Bruker EM27 with a weather station-controlled automated clamshell cover and for the solar tracker and protective fairing (Fig. 2). Our objective was to achieve autonomous and remote measurements of greenhouse gases in harsh environments, in particular, to perform much needed measurements of greenhouse gas (and) total columns X_{CO₂} and X_{CH₄}, in desert Australia. The design and more details of this construction can be obtained by contacting the authors.

Recently, Hedelius et al. (2016) published a long-term assessment of errors and biases in retrievals of several TCCON gases from the EM27. In spite of a reported drift, they found that the stability of the EM27 is sufficient for short-term campaigns. Therefore, before deployment to Alice Springs, we have operated the EM27 from December 2015 to September 2016 and tested the automated clamshell cover at the University of Wollongong, where the TCCON instrument (Griffith et al., 2014) is also located. Results of the collocated measurements are in Section 4.

3.2 Retrieval of X_{CO₂} and ~~X_{air}~~ X_{air}

To retrieve X_{CO₂}, we use the same retrieval software used by TCCON called GGG2014 (Wunch et al., 2017), which was also used by Hedelius et al. (2016) for their EM27. Column-averaged dry-air mole fractions (DMFs) of gases (X_{gas}) are retrieved from the EM27 measurement as in Wunch et al. (2010):

$$X_{gas} = \frac{VC_{gas}}{VC_{dryair}} = 0.2095 * \frac{VC_{gas}}{VC_{O_2}} \quad (1)$$

where VC_{gas} is the vertical column of the gas and VC_{dryair} is the dry pressure column of air and 0.2095 is the known DMF of oxygen. The DMF of air ("~~X_{air}~~X_{air}") is also retrieved using the oxygen column but uses the measured surface pressure to obtain the vertical column of air can be calculated using the measured oxygen column from the EM27 spectrum:

$$X_{air} = \frac{VC_{air}}{VC_{O_2}} * 0.2095 - X_{H_2O} * \frac{m_{H_2O}}{m_{air}^{dry}} \quad (2)$$

where m_{H_2O} and m_{air}^{dry} are the mean molecular masses of water (18.02 g mol⁻¹) and dry air (28.964 g mol⁻¹), X_{H_2O} is the retrieved DMF of water vapour and VC_{air} is calculated from the surface pressure, P_s :

$$VC_{air} = \frac{P_s}{g * \frac{m_{air}^{dry}}{N_a}} \quad (3)$$

where g is the column-averaged acceleration due to gravity and N_a is Avogadro's constant. X_{air} is a good indicator of instrument stability and changes in spectrometer alignment because 1) VC_{air} is calculated using the surface pressure, which ~~can be measured~~ is independently measured by a pressure sensor to better than 0.3 hPa (Wunch et al., 2011) ~~and accurate~~, keeping accuracy over long periods ~~with current instrumentation~~ (Wunch et al., 2011); and 2) the atmospheric oxygen column is not particularly variable in dry air, hence the retrieved VC_{O_2} by the spectrometer should be close to constant. From Eq. (2) it follows that a perfectly accurate measurement would lead to an X_{air} value of unity, however due to inaccuracies in the O_2 spectroscopy, the actual value is approximately 0.98 for all TCCON sites. We also find that X_{air} values well outside of the 2nd and 98th percentiles may indicate an obstruction in the solar beam (e.g. birds, fast moving clouds, leaves from trees, etc.). Therefore the X_{air} can also serve as data filtering criteria.

In Wollongong, both the TCCON and EM27 solar tracker covers were opened and closed by the same pneumatic system. The same weather station provided the meteorological data that were used to pre-filter the data (e.g. fractional variation in solar intensity, wind speed and direction, pressure, etc). For this study, we did not filter the data according to solar zenith angles (SZAs) anymore, instead in addition to the pre-filter, we filtered out noisy retrievals by selecting only those with X_{air} values within 0.5 and 1.5 because anything beyond that would be unrealistic in the atmosphere but most likely be the cause of an interference or obstruction.

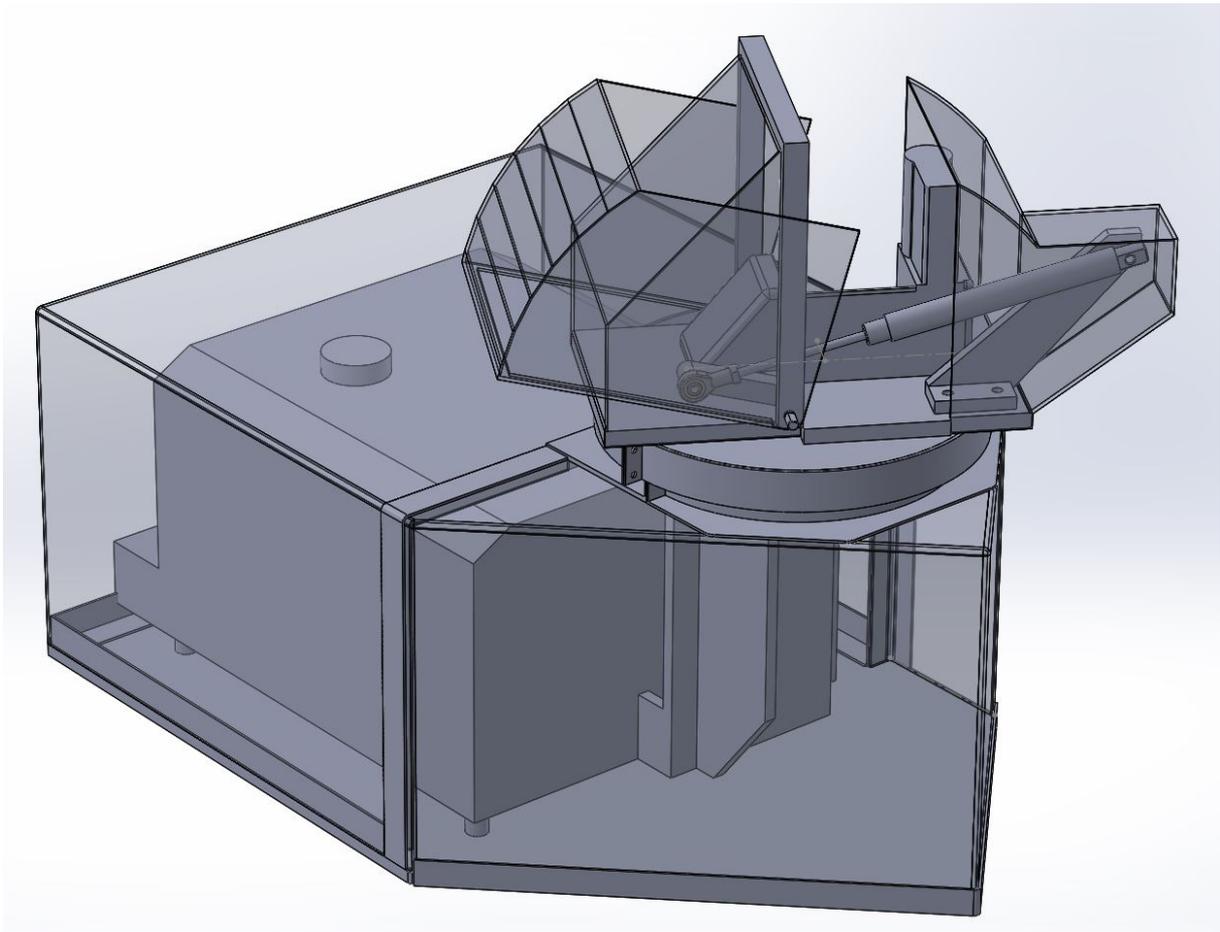


Figure 2. Drawing of the Wollongong EM27 automated clamshell cover and fairing (by Steve Selby, UOW).

3.3 GOSAT Specific Point Observations

The TANSO-FTS on GOSAT has a two-axis pointing system. Normally, TANSO-FTS follows an M shaped grid on a 5-point cross-track scan mode (https://www.eorc.jaxa.jp/GOSAT/instrument_1.html). By using this pointing system to vary the observation geometry, it is able to observe specific points, i.e. it can view targets with angles up to $\pm 20^\circ$ along the satellite track and by $\pm 30^\circ$ across the track. Specific point observations over Alice Springs were requested from July 2016, in preparation for the campaign in September. Five locations within 100 km of the center of Alice Springs were targeted (see Fig. 3). We only used the specific point observation data to compare with the EM27. However, to construct a the time-series shown in Fig. 78, we used all available GOSAT data version V2.72 from NIES spanning the years 2010 to 2017. Daily averages within 1000 km of Alice Springs are calculated for M- and H-gain retrievals separately.

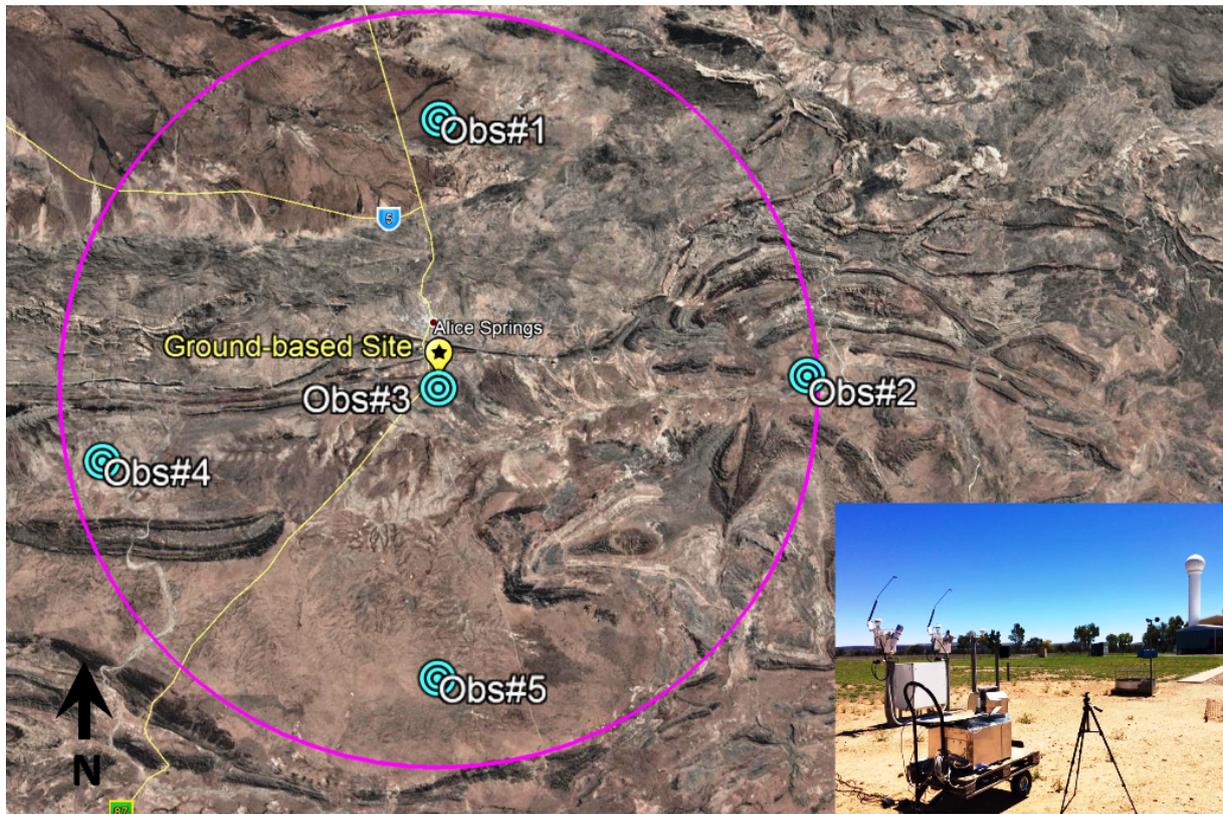


Figure 3. A Google Earth map showing the GOSAT specific point observations around Alice Springs (cyan circles labeled "Obs#") . The ground-based site is [securely fenced and](#) located at the airport, [eac](#). 14 km south of the town center. The magenta circle shows a 60-km radius from the site, for visual reference. [The inset shows the EM27 during the measurements.](#)

4 Results

4.1 Comparison with Wollongong TCCON Station

In this subsection, we present a side-by-side comparison of retrievals of X_{CO_2} between the EM27 and the Wollongong TCCON station. Apart from the [work by Hedelius et al. \(2016\)](#) [works by Hedelius et al. \(2016\)](#) and [Frey et al. \(2019\)](#), studies on long-term comparisons of the EM27 with TCCON are rare. Here, we focus on comparisons of $\overline{X_{air}} - X_{air}$ and X_{CO_2} from measurements spanning almost one year ([Nov. 2015 to Sep. 2016](#)) under varied environmental conditions. [Both instruments normally measure at the same time, apart from interruptions due to occasional mid-infrared measurements with the 125HR or rare software glitches \(e.g. JAVA issues\).](#)

4.1.1 $\overline{X_{air}}$

4.1.1 X_{air}

From Equation 2, we can see that the theoretical value of $\overline{X_{air}}$ would be 1.0 for a perfect instrument with a perfect retrieval. However, As mentioned in Sect. 3.2, inaccuracies in the Θ_2 - O_2 spectroscopy lead to an $\overline{X_{air}}$ - X_{air} value of approximately 0.98 for all TC-CON sites. This value varies by approximately 1% with solar zenith angle, indicating an airmass dependence in the Θ_2 - O_2 retrievals (Pollard et al., 2017). Deviations from the characteristic values for $\overline{X_{air}}$ - X_{air} are generally indicative of erroneous behavior in the measurement and retrieval system such as interferometer misalignment, tracking errors or fitting to an incorrect airmass due to timing errors, as recently reported in Pollard et al. (2017).

Figure 4 shows the normalized probability distribution functions (PDF) of $\overline{X_{air}}$ - X_{air} values retrieved by both instruments in 2016. One manifestation of instrumental differences between the Wollongong TCCON and the EM27 is the difference in the $\overline{X_{air}}$ - X_{air} retrieved from both instruments. The corresponding mean values of $\overline{X_{air}}$ - X_{air} from both instruments are 0.9824 and 0.9849 for TCCON and EM27 respectively. The TCCON $\overline{X_{air}}$ - X_{air} PDF exhibits a more Gaussian pattern, while the EM27 $\overline{X_{air}}$ - X_{air} PDF is slightly skewed.

Figure 5 shows the mean $\overline{X_{air}}$ - X_{air} values calculated for each month in 2016 for TCCON (red) and EM27 (blue). The error bars represent the spread (one standard deviation) in the monthly mean values of $\overline{X_{air}}$ - X_{air} . This should give us a measure of how much the $\overline{X_{air}}$ - X_{air} changes and therefore another measure of instrument stability. In Fig. 5, we see that the mean $\overline{X_{air}}$ - X_{air} values from the EM27 are slightly larger compared to the TCCON instrument but a slight seasonal dependence can be seen in both instruments. The difference in the mean $\overline{X_{air}}$ - X_{air} values does not appear to be purely related to sampling differently across that seasonal dependence. The SZAs seem to have an effect as well, although small. To show this in Fig. 5, we filtered X_{air} values for $SZA \leq 45^\circ$ (black dots and circles) and above 45° (grey pluses and crosses). The SZA effect is within approximately 1% for both instruments. We note that there were no re-alignments done on the EM27 and TCCON instruments during this period, however the EM27 clamshell cover and fairing were fitted on the EM27 in February 2016. It is possible that this may have affected the alignment resulting in a shift in the X_{air} .

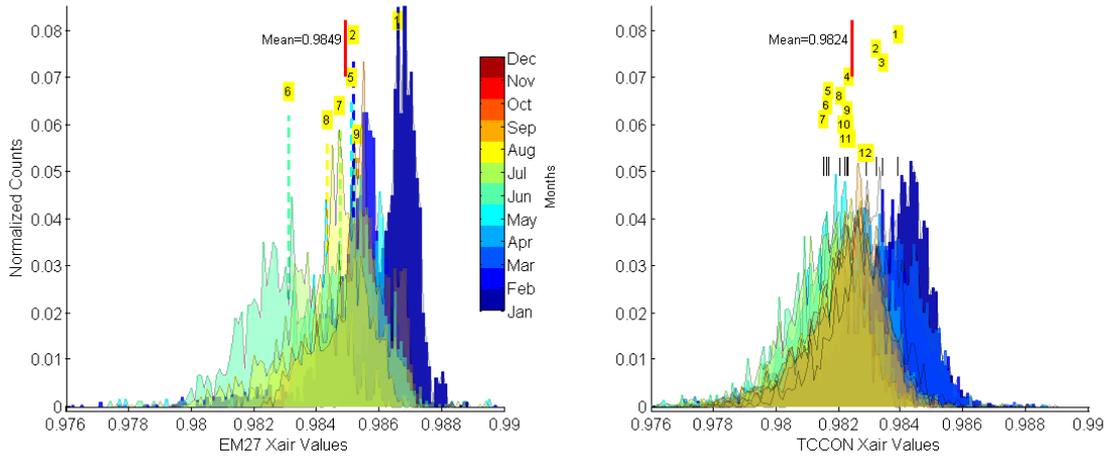


Figure 4. Normalized probability distribution functions (PDF) of X_{air} from TCCON (right panel) and X_{air} from the EM27 (left panel) in Wollongong for 2016. The PDFs for each month are separated by colors and the positions of the mean X_{air} - X_{air} monthly values for each month are indicated by the month number in a year. The mean X_{air} - X_{air} for the whole year for the EM27 is 0.9849 and for TCCON is 0.9824 for the whole year TCCON. The clamshell cover and fairing were installed on the EM27 in February, this may have affected the alignment resulting in a shift in the X_{air} .

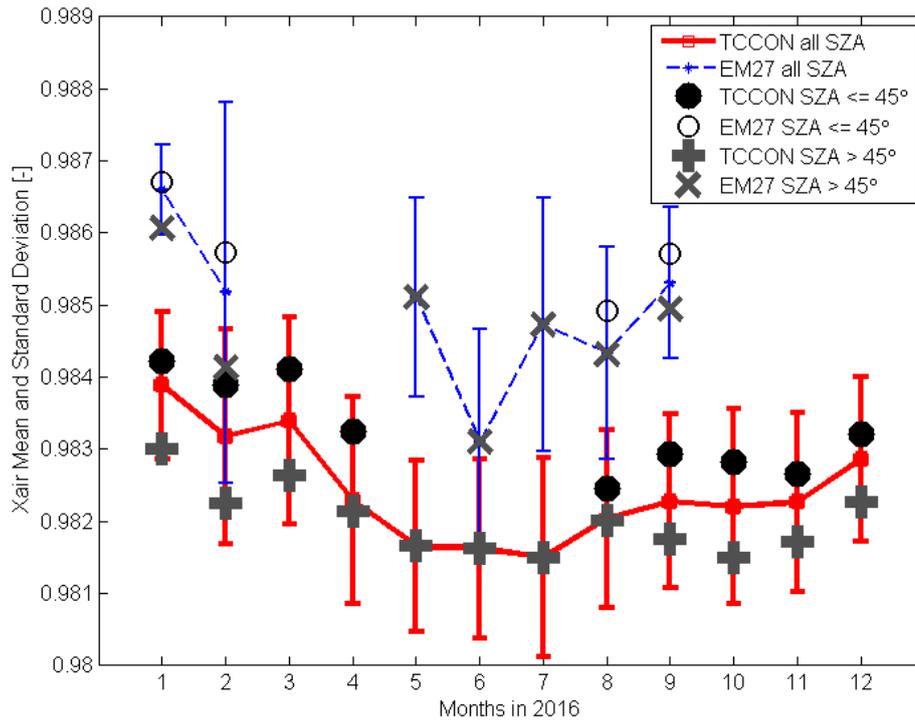


Figure 5. A measure of the spread of the of X_{air} values from TCCON (red) and EM27 (blue) by taking the mean (central markers) and corresponding standard deviations (bars) for each of the monthly probability distribution functions shown in Fig. 4. Filtered X_{air} values for angles $\leq 45^\circ$ (black dots and circles) and above 45° (grey pluses and crosses) are also shown for reference.

4.1.2 X_{CO_2} Comparison with Wollongong TCCON and Apparent Drift

From the EM27 and TCCON measurements in 2016, we derived a calibration-scaling factor for the EM27. We calibrated-scaled the EM27 measurements to Wollongong TCCON data because the Wollongong X_{CO_2} has been calibrated against aircraft profile measurements that are traceable to the WMO in-situ scale (Wunch et al., 2010). The retrieval method is predicted to be both linear and have zero intercept (Wunch et al., 2010). Therefore, we fit hourly mean data from TCCON and EM27 employing linear least squares and force a zero intercept. The standard errors of the weighted means are used as weights in the fit. From this exercise, we arrive at a calibration-scaling factor of $EM27=0.9954 * TCCON$ (Fig. 6). Frey et al. (2015) reported a comparable calibration-scaling factor of $0.9951 * TCCON$ for X_{CO_2} using collocated measurements of five EM27 instruments with the TCCON instrument at the Karlsruhe Institute of Technology in Germany. However, Hedelius et al. (2016) reported a smaller bias of $+0.03\%$ between EM27 and TCCON. We think that the slightly larger bias in this work (-0.46%) is probably due to an imperfect alignment of our EM27, with a modulation efficiency (ME) of 96% at maximum optical path difference (OPD). The ME calculation was done using water lines measured at ambient laboratory/room air (Frey et al., 2015). The

ILS was monitored before, during and after the campaign, and the ME remained stable within 94%-96%. Therefore, we are confident that the calibration scaling factor relative to TCCON was consistent during the campaign ~~with that derived relative to Wollongong TCCON~~.

5 Hedelius et al. (2016) reported a noticeable drift in the EM27 measurements over several months. We also observed a very small drift in the EM27 vs TCCON X_{CO_2} values shown in Fig. 6, where the calibration line has been derived. For clarity, we color-coded the EM27 vs TCCON X_{CO_2} values in Fig. 6 according to days after Jan 1, 2016. The colors progress with time, i.e. from dark blue (oldest) to dark brown (most recent). We found that the drift within three months before the campaign is very small, i.e. ~~0.0125% per month~~ 0.13% over three months as calculated from the EM27/TCCON calibration scaling factors from July to mid September. We did not consider the averaging kernels in this work. The averaging kernels of the EM27 have been
10 previously presented and compared to TCCON in a study by Hedelius et al. (2016). In their study, they found that although there are differences in the TCCON and EM27 averaging kernels, the effect of the differences in averaging kernels from the top of the atmosphere cancel out the effect of differences at the bottom. Further work and more measurements may be necessary to better understand the cause of this phenomenon but this is beyond the scope of this study.

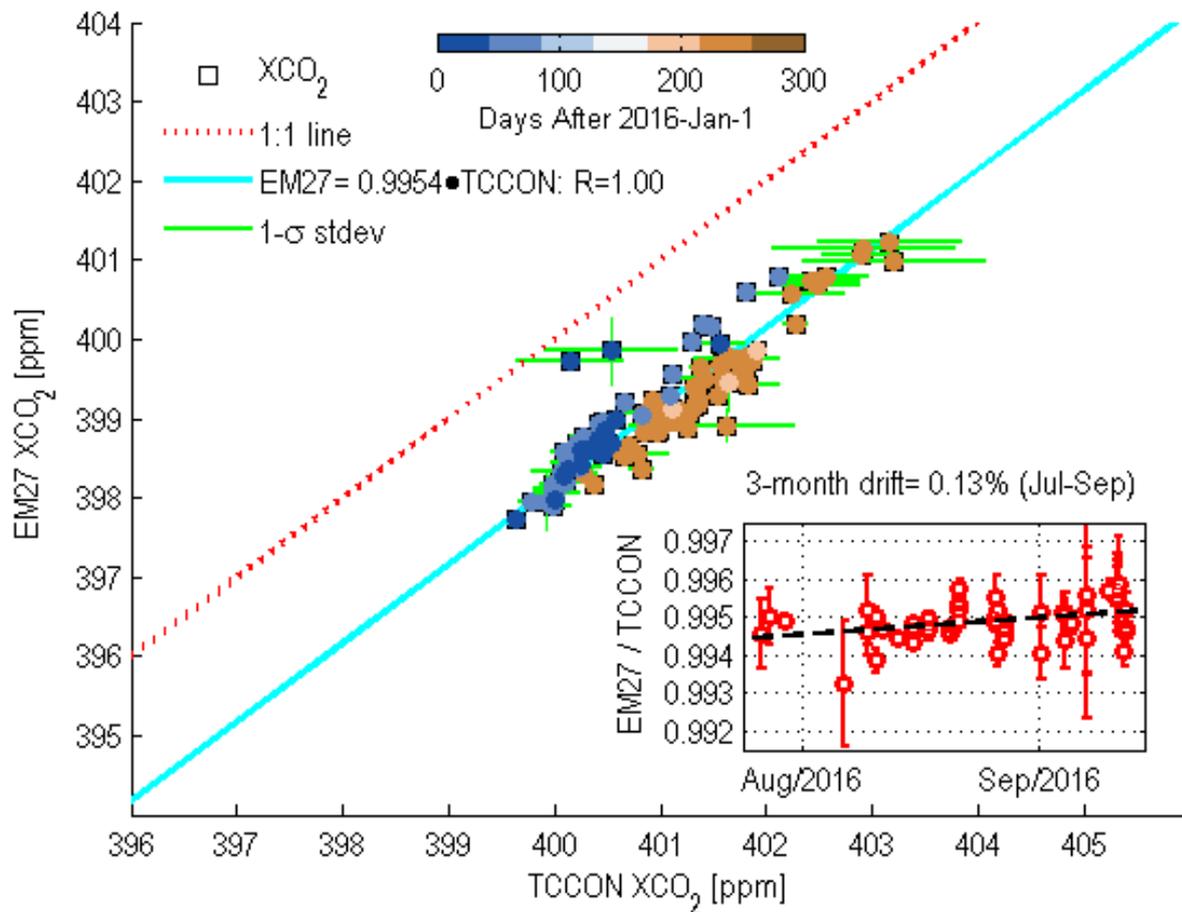


Figure 6. Retrievals of Weighted hourly mean X_{CO_2} from TCCON vs X_{CO_2} from EM27 in Wollongong for 2015-2016. To avoid noisy data, only measurements corresponding to TCCON with retrieved X_{air} values within 0.9783 and 0.9853 were selected (X_{air} values within the 2nd to 98th percentile). The inset shows the drift in the hourly weighted mean scaling factors over three months (red dots). The error bars are the sum, in quadrature, of the relative standard deviations of the weighted hourly means from each instrument.

4.2 Measurements in Alice Springs

Results of the the ground based X_{CO_2} measurements during the Alice Springs campaign are shown in Fig. 7. The measurements started on September 29, 2016 with interruptions due to cloudy and/or rainy weather. In total, we gathered 6 days of good ground-based measurements of X_{CO_2} under good weather conditions, which coincided with 9 GOSAT specific point soundings within 100 km of the site. A slight airmass dependence in the X_{CO_2} retrievals results ~~to~~ in lower X_{CO_2} values at low sun elevation, as can be observed in Fig. 7. This airmass dependence is well-known and also discussed in Wunch et al. (2011). But for the purpose of GOSAT comparisons, this airmass dependence is negligible because GOSAT passes the site at around 13:05 local time (03:35 UTC), corresponding to a high sun elevation (see Table 1 for the dates).

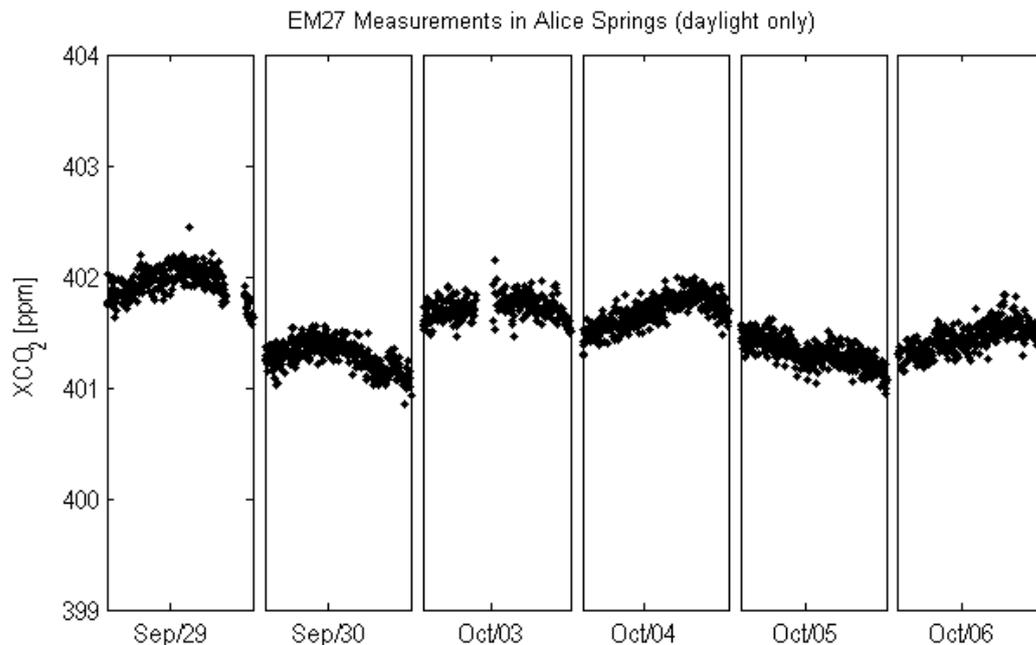


Figure 7. Retrievals of X_{CO_2} from the EM27 in Alice Springs, compressed along the x-axis (days). The measurement days on the x-axis is are not continuous due to interruption-interruptions from bad weather. See Section-Sec_6 for data availability.

4.3 Comparisons with GOSAT Measurements in Alice Springs

To demonstrate the variability of X_{CO_2} in the region, we plot ~~the time series of interpolated GOSAT soundings near Alice Springs~~ in Fig. 8. ~~The daily mean soundings were selected and calculated~~ the daily mean time series of GOSAT soundings using these two coincidence criteria: 1) all soundings within a 1000-km radius of-centered at the BOM facility where EM27
 5 ~~measurements were made in Alice Springs and 2) all soundings measured within the same day (local time). Each data point considered in the calculation of the daily mean is weighted by the corresponding reported retrieval error for that particular GOSAT sounding.~~ Vertical gray lines represent the corresponding standard deviation for each daily mean calculation. The H-gain retrievals (red triangles) and M-gain retrievals (blue squares) ~~were~~ are separated in this plot. Weighted linear least squares fitted straight lines were calculated for the M-gain X_{CO_2} (cyan) and H-gain X_{CO_2} data (magenta). We confirm that the M-gain
 10 retrievals are biased high (around 2 ppm) compared to the H-gain retrievals. Nevertheless, both least squares fitted lines to the H- and M-gain retrievals show an increase of about 2.28 ppm/year with y-intercept values at 133.0 ppm (H-gain) and 134.92 ppm (M-gain).

Compared to the TCCON sites in Wollongong and Darwin (~~Deutscher et al., 2014~~)(Deutscher et al., 2010, 2014), the X_{CO_2} signal in Alice Springs is relatively smooth and undisturbed, which is expected because Alice Springs is in the middle of the
 15 Australian continent, a desert environment with no large sources to interfere with the X_{CO_2} signal. The terrestrial biosphere

is the largest driver of variability in the Southern Hemisphere column X_{CO_2} (Deutscher et al., 2014). Moreover, Deutscher et al. (2014) have shown that the magnitude of the seasonal variability in the column-average dry-air mole fraction of X_{CO_2} is comparable in magnitude to the annual increase. In Alice Springs, only a slight seasonal cycle in the measured X_{CO_2} can be seen. This slight cycle is probably less driven by the surrounding vegetation because the centre of the Australian arid zone is not affected by seasonality and has low aggregate vegetation (Lawley et al., 2011) but more by meridional effect.

~~However, H-gain retrievals seem to have diminished around 2011, coinciding with the dry months starting from April 2011. H-gain soundings from Rainfall records indicate that April-August 2011 were 23% less compared to the April-August mean derived from the years 2010-2017. This could be due to stunted vegetation growth or mortality in the had the least amount of rainfall on record in Alice Springs after 2002, which was only broken by 2015, then by 2017 (http://www.bom.gov.au/climate/data/station_number_015590). This significant absence of rain in the region, which could result in stunted vegetation growth or mortality, which could have led to brighter surfaces. Bright surfaces mean more M-gain measurements and this may explain the 74% increase in M-gain soundings in April-August 2011 compared to the April-August 2010-2017 average. Rainfall records indicate that of c. 220 soundings/month. In contrast, the number of H-gain retrievals seems to have diminished around 2011, coinciding with the dry months starting from April 2011. The number of H-gain soundings from April-August 2011 had the least amount of rainfall on record in Alice Springs after 2002, which was only broken by 2015, then 2017 (http://www.bom.gov.au/climate/data/station_number_015590 month).~~

GOSAT specific point observations that were scheduled during the campaign are shown on Table 1. ~~M and~~ The times correspond to the times when the satellite is directly above the site. The satellite normally performs 5 observation points across track, with an interferometric scan time lasting 4 seconds (Shiomi et al., 2006). But specific point observations deviate from this pattern by pointing and maximizing observations near the target. M- and H-gain retrievals (6 days apart) were obtained during the short campaign. We averaged and compared the GOSAT retrievals directly with the coincident measurements from the EM27 (± 0.5 hour) and ~~derived a calibration as in Sec. 4.1.2, we fit the data with using weighted linear least squares and force a zero intercept. From here, we derived a scaling~~ factor of $0.9927 \cdot EM27$ for GOSAT H-gain and $0.9983 \cdot EM27$ for the GOSAT M-gain, as shown in Fig. 9. This is only slightly different from the previous version 2.60 of the data (H-gain: $0.9935 \cdot EM27$ and M-gain: $0.9997 \cdot EM27$). Indeed, there is a slight bias in the ~~M and H-gain~~ M- and H-gain retrievals from GOSAT, i.e. ~~M-gain~~ M-gain retrievals are biased slightly low and H-gain retrievals are very close compared to the EM27. However, the ~~M-M-~~ M- and H-gain bias relative to each other has been improved in the version 2.72 data release.

Table 1. GOSAT Specific Point Observation Opportunities Over Alice Springs

Date (UTC)	Orbit Number	Gain Type (H or M)	Remarks
2016/09/27 03:35 :28	4	-	soundings > 100 km
2016/09/30 03:35 :28	4	H	4 soundings
2016/10/03 03:35 :28	4	-	soundings > 100 km
2016/10/06 03:35 :28	4	M	5 soundings

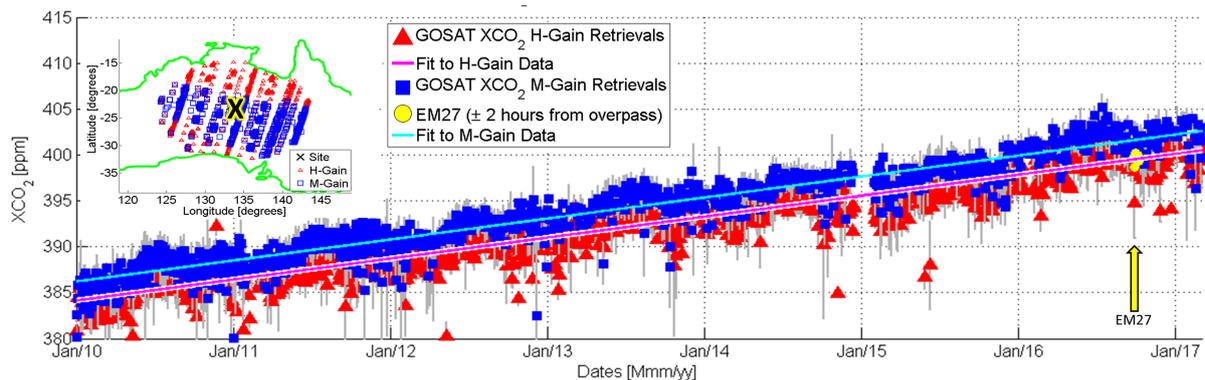


Figure 8. Time series X_{CO_2} from GOSAT M-gain and H-gain retrievals in Alice Springs from the NIES version 2.72 product. The timing of the Alice Springs campaign from Sep. 27 to Oct. 6 is indicated by the yellow arrow. Gray lines represent one standard deviation error associated with each daily mean. The inset shows a map projection with the approximate locations of the soundings used for this plot. Only data within a 1000-km radius from Alice Springs (yellow-black X on the map) were considered.

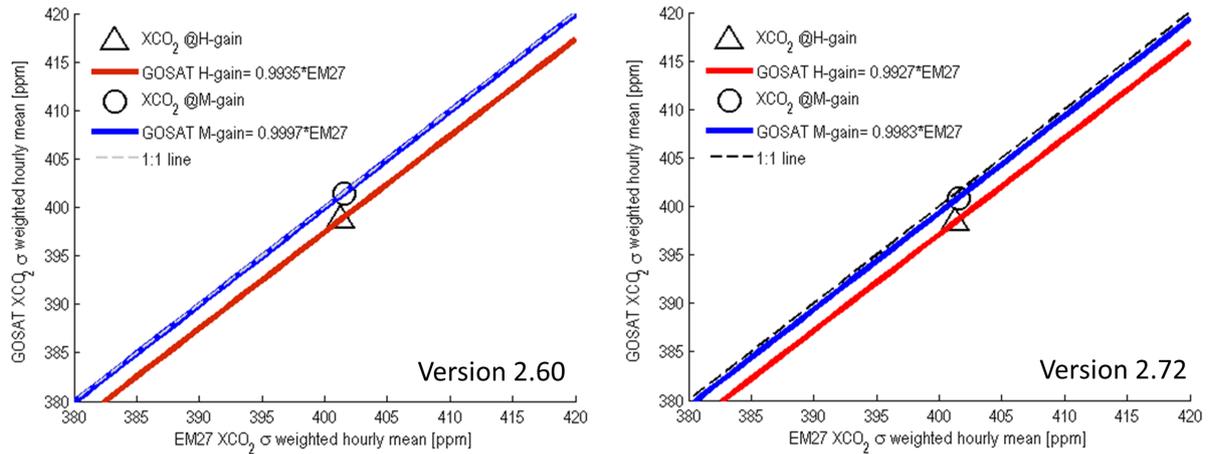


Figure 9. Mean X_{CO_2} from the EM27 vs GOSAT X_{CO_2} version 2.60 (left panel) and version 2.72 (right panel) in Alice Springs. The calculated means are weighted by the $1-\sigma$ uncertainty of the individual measurements. The scaling factor derived from the TCCON comparisons, $EM27=0.9954*TCCON$, has been taken into account.

4.4 Recommendations for a Future Alice Springs Campaign: Simulating Error Reduction in Bias Estimates

In this subsection, we attempt to estimate how long a measurement campaign in Alice Springs should last in order to improve the statistical errors in the estimated bias between GOSAT and the EM27. In order to estimate the required measurements from the EM27, we make the following assumptions and conditions:

- 5 1. At least one specific point overpass per week (M-gain or H-gain), similar to this campaign.
2. ~~1- σ standard deviation of all~~ The standard errors of the weighted hourly mean GOSAT specific point observations ~~is~~ are: 0.25 ppm (~~M-gain~~ and 0.26 ppm for M-gains (versions 2.6 and 2.72, respectively)) and 0.29 ppm (~~H-gain~~ and 0.28 ppm for H-Gains (versions 2.6 and 2.72, respectively)). These values are taken from the averaged standard errors of the weighted hourly means from all specific point observation data falling within 100 km from the site for 2 Sep. 2015 - Feb. 2017.
- 10 3. $1-\sigma$ standard deviation of all EM27 measurements is 0.05 ppm at an averaging time of 1-hour. Chen et al. (2016) reported a $1-\sigma$ standard deviation of 0.04 ppm at 10 minutes averaging.
4. The error in the estimated bias is calculated as $\sigma_T = \sqrt{\sigma_{GOSAT}^2 + \sigma_{EM27}^2}$, which then improves by the number of weeks (N), i.e. $\sqrt{\sigma_T^2/N}$
- 15 5. Any long-term drifts in the satellite and ground-based measurements are negligible over the whole duration of the campaign.

It is important to keep the statistical errors in the estimated bias as low as possible. Miller (2007) showed that a comparison of surface CO_2 concentration data and X_{CO_2} data flux inversions clearly reveals a land-ocean bias in the X_{CO_2} retrievals, even when the bias is only 0.1 ppm. Fig. 10 shows a plot of how the statistical errors in the estimated bias improve with the number of weeks in the campaign. The inset shows the normalized histogram of the reported GOSAT single sounding errors within 100 km of the site. Note that although H-gain retrievals have smaller errors on average, the amount of M-gain retrievals improves the standard error of the weighted mean. A two-week campaign would already reduce the statistical error between GOSAT M-gain and EM27 measurements to less than the reported GOSAT M-Gain standard deviation (0.25 ppm weighted mean standard error (0.26 ppm, v2.72)). To achieve a statistical error of 0.1 ppm between GOSAT M-gain and EM27 measurements, it would take about 6.5-6 to 7 weeks of measurements or more. On the other hand, around 8.5 weeks for H-gain is needed because the standard error of the hourly weighted mean H-gain retrievals have a slightly higher standard deviation are slightly higher due to fewer soundings near the site. The ideal time window to perform this would be from March to November (Autumn-Winter-Spring). It is best to avoid the summer months (Dec-Feb) because of the desert heat in Alice Springs, which could be physically demanding for the operator/s and may affect the operation of an EM27 without a temperature-regulated cover enclosure. A temperature-regulated cover enclosure would improve stability and make the EM27 suitable for summer operations even in the high heat ($>40^\circ\text{C}$) of Alice Springs. Note however that our error estimates are conservative and this implies the assumption that there are no drifts in the M- and H-gain measurements relative to one another that would require long-term validation.

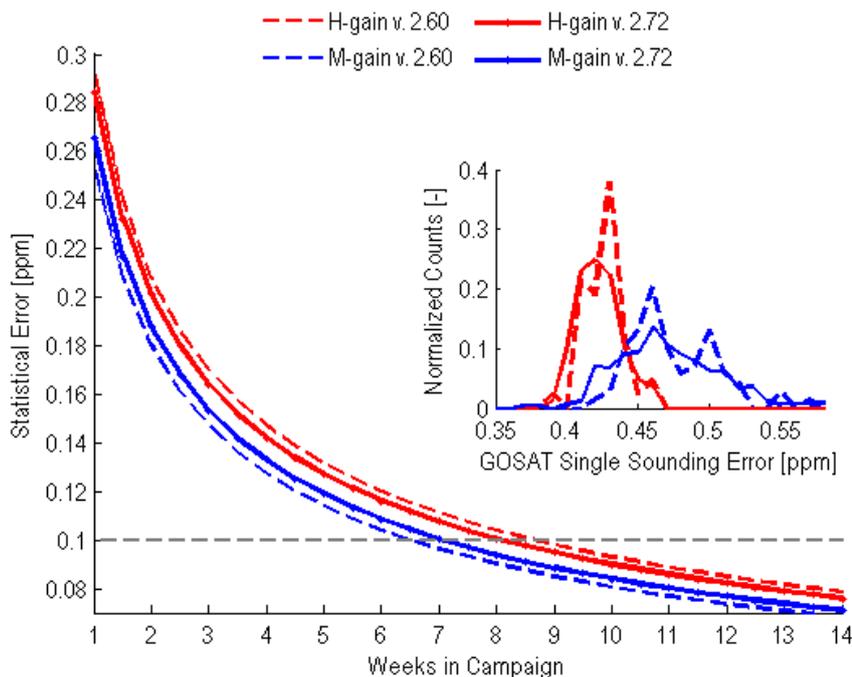


Figure 10. The statistical errors in the estimated bias of X_{CO_2} improve with the number of weeks in the campaign. Inset: normalized histogram of the reported GOSAT single sounding errors within 100 km of the site.

5 Conclusions

Alice Springs would be a welcome addition to ground sites dedicated to environmental satellite sensor calibration and validation. With the exception of Lauder, New Zealand, there are no permanently dedicated ground targets for satellite calibration and validation in the Southern Hemisphere. ~~Any addition from the southern region would be highly valued, especially the~~
5 ~~type of addition that Alice Springs could offer~~that are far enough from water (to avoid a possible land-ocean bias in retrievals)
and surrounded by homogeneous topography (to avoid a possible altitude bias). Currently, vicarious radiometric calibration of passive optical instruments from satellites is done on the large homogeneous Railroad Valley playa (RRV) in Nevada, USA. However plans for future lithium mining threaten to disrupt the playa floor, which could render this critical asset useless for large footprint sensors on satellites like GOSAT, GOSAT-2~~and OCO-2, OCO-3 and on other future satellites~~, OCO-2 and
10 OCO-3. Alice Springs would offer unique opportunities particularly in the context of satellite surface products and associated emissivity assessments because the environment of Alice Springs ~~provides~~ has ideal conditions to measure undisturbed atmospheric conditions that ~~will~~ can serve as calibration point for satellite retrievals of atmospheric composition. At the same time, the desert environment, like RRV, provides high surface reflectivity, which is a challenge ~~to~~ for satellite retrievals and thus ~~this data is especially needed~~ observations are especially needed here for validation purposes. Subject to funding and interest from
15 the community, a repeat of the campaign is possible. In a future study, we will investigate GOSAT methane retrievals using the same approach employed here.

6 Data Availability

Column averaged dry-air mole fractions of CO₂ in Alice Springs, Australia is hosted by the University of Wollongong and can be downloaded via <http://ro.uow.edu.au/data/58/> with the DOI: <http://dx.doi.org/10.4225/48/5b21f16ce69bc> (Velazco et al.,
20 2018). GOSAT data can be obtained from the GOSAT Data Archive Service (GDAS) after registration for access via: https://data2.gosat.nies.go.jp/index_en.html.

Author contributions.

V.V. conceptualized and wrote the manuscript, did the calculations and statistics, made the figures, performed EM27 measurements in Alice Springs, led the EM27 automated clamshell cover project, ~~calculated~~ characterized the EM27 and TCCON
25 ILS and oversaw TCCON measurements in Wollongong during this study.

N.D. helped in writing the manuscript, organized and ~~co-funded~~ co-acquired funding for the Alice Springs campaign, liaised with BOM, performed the GGG2014 retrievals of the EM27 data, performed measurements in Alice Springs, oversaw the EM27 measurements in Wollongong.

I.M. helped in writing the manuscript, provided ideas, conceptualized the need for statistical calculations (section 3-4.4).

O.U. presented and highlighted the M-M- and H-gain bias and the need for validating GOSAT in M-gain region as representative of the GOSAT project in the RA-Research Announcement (RA) meeting, helped in writing the manuscript and provided adviseadvice.

5 D.G. supervised the EM27 automated clamshell cover project, supervised the EM27 and TCCON measurements, created the preprocessing software, helped in writing the manuscript and provided advise.

N.J. carried out EM27 measurements in Wollongong and helped in determining the EM27 ILS.

C.P.W. performed TCCON measurements in Wollongong and helped improve the manuscript.

B.B. carried out TCCON measurements in Wollongong and helped improve the manuscript.

A.K. and M.A. enabled the specific observation mode in Alice Springs from 2016.

10 *Competing interests.* The authors declare that they have no conflict of interest.

Acknowledgements. V.V. would like to thank University of Wollongong SMAH for the advancement grants scheme that funded the fabrication of the EM27 clamshell cover and fairing (Grant number 2016/SPGA-S/07). We would like to thank the Australian Bureau of Meteorology (BOM), especially site manager Victoria McLean for the support in Alice Springs. We thank Graham Kettlewell for technical advise, programming the weather station and liaising with the workshop. We thank Martin Rikkenbach for the design and creation of the weather station
15 interface and for providing technical advise. We thank Steve Selby for the design and fabrication of the EM27 clamshell and cover and Phd student Neil Page for assistance with the EM27 measurements. The specific point observation modes have been made possible by the GOSAT Project Office. Part of this activity has been conducted under the framework of the GOSAT Research Announcement 8 (RA8) Project. N.D. acknowledges funding from the Australian Research Council (ARC) via DECRA grant DE140100178. Wollongong TCCON activities have been funded through ARC Discovery Projects (DP0879468, DP110103118, DP140101552, DP160101598) and LIEF infrastructure grant
20 LE0668470.

References

- Ahlström, A., Raupach, M. R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J. G., Friedlingstein, P., Jain, A. K., Kato, E., Poulter, B., Sitch, S., Stocker, B. D., Viovy, N., Wang, Y. P., Wiltshire, A., Zaehle, S., and Zeng, N.: The dominant role of semi-arid ecosystems in the trend and variability of the land CO₂ sink, *Science*, 348, 895, <http://science.sciencemag.org/content/348/6237/895.abstract>, 2015.
- 5 Borsdorff, T., Aan De Brugh, J., Hu, H., Hasekamp, O., Sussmann, R., Rettinger, M., Hase, F., Gross, J., Schneider, M., Garcia, O., Stremme, W., Grutter, M., Feist, D. G., Arnold, S. G., De Mazière, M., Kumar Sha, M., Pollard, D. F., Kiel, M., Roehl, C., Wennberg, P. O., Toon, G. C., and Landgraf, J.: Mapping carbon monoxide pollution from space down to city scales with daily global coverage, *Atmospheric Measurement Techniques*, <https://doi.org/10.5194/amt-11-5507-2018>, 2018.
- 10 Chen, J., Viatte, C., Hedelius, J. K., Jones, T., Franklin, J. E., Parker, H., Gottlieb, E. W., Wennberg, P. O., Dubey, M. K., and Wofsy, S. C.: Differential column measurements using compact solar-tracking spectrometers, *Atmospheric Chemistry and Physics*, <https://doi.org/10.5194/acp-16-8479-2016>, 2016.
- Crisp, D., Pollock, H. R., Rosenberg, R., Chapsky, L., Lee, R. A. M., Oyafuso, F. A., Frankenberg, C., O'Dell, C. W., Bruegge, C. J., Doran, G. B., Eldering, A., Fisher, B. M., Fu, D., Gunson, M. R., Mandrake, L., Osterman, G. B., Schwandner, F. M., Sun, K., Taylor, T. E.,
- 15 Wennberg, P. O., and Wunch, D.: The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, *Atmospheric Measurement Techniques*, 10, 59–81, <https://doi.org/10.5194/amt-10-59-2017>, 2017.
- Deutscher, N. M., Griffith, D. W. T., Bryant, G. W., Wennberg, P. O., Toon, G. C., Washenfelder, R. A., Keppel-Aleks, G., Wunch, D., Yavin, Y., Allen, N. T., Blavier, J.-F., Jiménez, R., Daube, B. C., Bright, A. V., Matross, D. M., Wofsy, S. C., and Park, S.: Total column CO₂ measurements at Darwin Australia: site description and calibration against in situ aircraft profiles, *Atmospheric Measurement Techniques*,
- 20 3, 947–958, <https://doi.org/10.5194/amt-3-947-2010>, 2010.
- Deutscher, N. M., Sherlock, V., Fletcher, S. E. M., Griffith, D. W. T., Notholt, J., Macatangay, R., Connor, B. J., Robinson, J., Shiona, H., Velazco, V. A., Wang, Y., Wennberg, P. O., and Wunch, D.: Drivers of column-average CO₂ variability at Southern Hemispheric Total Carbon Column Observing Network sites, *Atmospheric Chemistry and Physics*, 14, 9883–9901, <https://doi.org/10.5194/acp-14-9883-2014>, 2014.
- 25 Frey, M., Hase, F., Blumenstock, T., Groß, J., Kiel, M., Mengistu Tsidu, G., Schäfer, K., Sha, M. K., and Orphal, J.: Calibration and instrumental line shape characterization of a set of portable FTIR spectrometers for detecting greenhouse gas emissions, *Atmospheric Measurement Techniques*, <https://doi.org/10.5194/amt-8-3047-2015>, 2015.
- Frey, M., Sha, M. K., Hase, F., Kiel, M., Blumenstock, T., Harig, R., Surawicz, G., Deutscher, N. M., Shiomi, K., Franklin, J. E., Bösch, H., Chen, J., Grutter, M., Ohshima, H., Sun, Y., Butz, A., Mengistu Tsidu, G., Ene, D., Wunch, D., Cao, Z., Garcia, O., Ramonet, M.,
- 30 Vogel, F., and Orphal, J.: Building the Collaborative Carbon Column Observing Network (COCCON): Long-term stability and ensemble performance of the EM27/SUN Fourier transform spectrometer, *Atmospheric Measurement Techniques*, <https://doi.org/10.5194/amt-12-1513-2019>, 2019.
- Gisi, M., Hase, F., Dohe, S., Blumenstock, T., Simon, A., and Keens, A.: XCO₂-measurements with a tabletop FTS using solar absorption spectroscopy, *Atmospheric Measurement Techniques*, <https://doi.org/10.5194/amt-5-2969-2012>, 2012.
- 35 Griffith, D. W. T., Velazco, V. A., Deutscher, N. M., Murphy, C., Jones, N., Wilson, S., Macatangay, R., Kettlewell, G., Buchholz, R., and Riggenschach, M.: Total Column Carbon Observing Network (TCCON) Site: Wollongong, <https://doi.org/10.14291/tcon.ggg2014.wollongong01.R0/1149291>, 2014.

- Hase, F., Frey, M., Blumenstock, T., Groß, J., Kiel, M., Kohlhepp, R., Mengistu Tsidu, G., Schäfer, K., Sha, M. K., and Orphal, J.: Application of portable FTIR spectrometers for detecting greenhouse gas emissions of the major city Berlin, *Atmospheric Measurement Techniques*, <https://doi.org/10.5194/amt-8-3059-2015>, 2015.
- 5 Hedelius, J. K., Viatte, C., Wunch, D., Roehl, C. M., Toon, G. C., Chen, J., Jones, T., Wofsy, S. C., Franklin, J. E., Parker, H., Dubey, M. K., and Wennberg, P. O.: Assessment of errors and biases in retrievals of XCO₂, XCH₄, XCO, and XN₂O from a 0.5 cm⁻¹ resolution solar-viewing spectrometer, *Atmospheric Measurement Techniques*, <https://doi.org/10.5194/amt-9-3527-2016>, 2016.
- Houweling, S., Baker, D., Basu, S., Boesch, H., Butz, A., Chevallier, F., Deng, F., Dlugokencky, E. J., Feng, L., Ganshin, A., Hasekamp, O., Jones, D., Maksyutov, S., Marshall, J., Oda, T., O'Dell, C. W., Oshchepkov, S., Palmer, P. I., Peylin, P., Poussi, Z., Reum, F., Takagi, H., Yoshida, Y., and Zhuravlev, R.: An intercomparison of inverse models for estimating sources and sinks of CO₂ using GOSAT measurements, *Journal of Geophysical Research*, <https://doi.org/10.1002/2014JD022962>, 2015.
- 10 Iwasaki, C., Imasu, R., Bril, A., Oshchepkov, S., Yoshida, Y., Yokota, T., Zakharov, V., Griбанov, K., and Rokotyan, N.: Optimization of the Photon Path Length Probability Density Function-Simultaneous (PPDF-S) Method and Evaluation of CO₂ Retrieval Performance Under Dense Aerosol Conditions, *Sensors*, <https://doi.org/10.3390/s19051262>, 2019.
- Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, *Applied Optics*, 48, 6716, <https://doi.org/10.1364%2Fao.48.006716>, 2009.
- 15 Lawley, E., Lewis, M., and Ostendorf, B.: Environmental zonation across the Australian arid region based on long-term vegetation dynamics, *Journal of Arid Environments*, <https://doi.org/10.1016/j.jaridenv.2011.01.003>, 2011.
- Matsunaga, T., Morino, I., Yoshida, Y., Saito, M., Noda, H., Ohshima, H., Kamei, A., Kawazoe, F., Maksyutov, S., Yokota, T., Kuze, A., Suto, H., Nakajima, M., and Imasu, R.: FTS and FTS-2 Onboard Japanese GOSAT Earth Observation Satellite Series, p. FW2B.2, <https://doi.org/10.1364/FTS.2018.FW2B.2>, 2018.
- 20 Miller, C. E., e. a.: Precision requirements for space-based XCO₂ data, *Journal of Geophysical Research Atmospheres*, <https://doi.org/10.1029/2006JD007659>, 2007.
- Morino, I., Uchino, O., Inoue, M., Yoshida, Y., Yokota, T., Wennberg, P. O., Toon, G. C., Wunch, D., Roehl, C. M., Notholt, J., Warneke, T., Messerschmidt, J., Griffith, D. W., Deutscher, N. M., Sherlock, V., Connor, B., Robinson, J., Sussmann, R., and Rettinger, M.: Preliminary validation of column-averaged volume mixing ratios of carbon dioxide and methane retrieved from GOSAT short-wavelength infrared spectra, *Atmospheric Measurement Techniques*, <https://doi.org/10.5194/amt-4-1061-2011>, 2011.
- 25 Parker, R., Boesch, H., Byckling, K., Webb, A., Palmer, P., Feng, L., Bergamaschi, P., Chevallier, F., Notholt, J., Deutscher, N., Warneke, T., Hase, F., Sussmann, R., Kawakami, S., Kivi, R., Griffith, D., and Velasco, V.: Assessing 5 years of GOSAT Proxy XCH₄ data and associated uncertainties, *Atmospheric Measurement Techniques*, 8, <https://doi.org/10.5194/amt-8-4785-2015>, 2015.
- 30 Pollard, D. F., Sherlock, V., Robinson, J., Deutscher, N. M., Connor, B., and Shiona, H.: The Total Carbon Column Observing Network site description for Lauder, New Zealand, *Earth System Science Data*, <https://doi.org/10.5194/essd-9-977-2017>, 2017.
- Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., Broquet, G., Canadell, J. G., Chevallier, F., Liu, Y. Y., Running, S. W., Sitch, S., and van der Werf, G. R.: Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle, *Nature*, 509, 600, <http://dx.doi.org/10.1038/nature13376>, 2014.
- 35 Reuter, M., Buchwitz, M., Hilker, M., Heymann, J., Schneising, O., Pillai, D., Bovensmann, H., Burrows, J. P., Bösch, H., Parker, R., Butz, A., Hasekamp, O., O'Dell, C. W., Yoshida, Y., Gerbig, C., Nehrkorn, T., Deutscher, N. M., Warneke, T., Notholt, J., Hase, F., Kivi,

- R., Sussmann, R., Machida, T., Matsueda, H., and Sawa, Y.: Satellite-inferred European carbon sink larger than expected, *Atmospheric Chemistry and Physics*, <https://doi.org/10.5194/acp-14-13739-2014>, 2014.
- Shiomi, K., Yoshida, M., and Sekio, N.: Study for sun glint observation of GOSAT FTS using MODIS and AMSR-E data, in: *International Geoscience and Remote Sensing Symposium (IGARSS)*, <https://doi.org/10.1109/IGARSS.2006.533>, 2006.
- 5 Suto, H., Yoshida, J., Desbiens, R., Kawashima, T., and Kuze, A.: Characterization and correction of spectral distortions induced by microvibrations onboard the GOSAT Fourier transform spectrometer., *Applied optics*, <https://doi.org/10.1364/AO.52.004969>, 2013.
- Trudinger, C. M., Haverd, V., Briggs, P. R., and Canadell, J. G.: Interannual variability in Australia's terrestrial carbon cycle constrained by multiple observation types, *Biogeosciences*, 13, 6363–6383, <https://doi.org/10.5194/bg-13-6363-2016>, <https://www.biogeosciences.net/13/6363/2016/>, 2016.
- 10 Turner, A., Jacob, D., Wecht, K., Maasackers, J., Lundgren, E., Andrews, A., Biraud, S., Boesch, H., Bowman, K., Deutscher, N., Dubey, M., Griffith, D., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker, R., Payne, V., Sussmann, R., Sweeney, C., Velazco, V., Warneke, T., Wennberg, P., and Wunch, D.: Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, *Atmospheric Chemistry and Physics*, 15, <https://doi.org/10.5194/acp-15-7049-2015>, 2015.
- Velazco, V. A., Deutscher, N. M., and Griffith, D. W. T.: Column averaged dry mole fractions of carbon dioxide in Alice Springs, Australia, <https://doi.org/http://dx.doi.org/10.4225/48/5b21f16ce69bc>, <https://ro.uow.edu.au/data/58/>, 2018.
- 15 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., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo,
- 20 M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, *Atmospheric Measurement Techniques*, 3, 1351–1362, <https://doi.org/10.5194%2Famt-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, *Philosophical Transactions of the Royal Society A: Mathematical Physical and Engineering Sciences*, 369, 2087–2112, <https://doi.org/10.1098%2Frsta.2010.0240>, 2011.
- 25 Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, C., Feist, D. G., and Wennberg, P. O.: Documentation for the 2014 TCCON Data Release, <https://doi.org/10.14291/TCCON.GGG2014.DOCUMENTATION.R0/1221662>, <https://tcccon-wiki.caltech.edu/>, 2017.
- Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global Concentrations of CO₂ and CH₄ Retrieved from GOSAT: First Preliminary Results, *SOLA*, 5, 160–163, <https://doi.org/10.2151/sola.2009-041>, 2009.
- Yoshida, Y., Kikuchi, N., Morino, I., Uchino, O., Oshchepkov, S., Bril, A., Saeki, T., Schutgens, N., Toon, G. C., Wunch, D., Roehl, C. M.,
- 30 Wennberg, P. O., Griffith, D. W., 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*, <https://doi.org/10.5194/amt-6-1533-2013>, 2013.