Response to reviewer 2

The authors appreciate the comments and corrections of the reviewers.

Specific comments:

The optical filter used for this study differ from the NDACC standard filter set. NDACC filter #1 covers the region from 3850–4550 cm-1 and does not include the 4800 cm-1 micro window. The filter used in this study is optional and in NDACC numbering (increasing number with increasing wavelength) something like NDACC filter #0. Therefore, it is questionable whether many sites recorded spectra with it since a long time.
Indeed is not a guarantee that this filter is being used in other sites. Ny-Alesund has one of the longest time series in the network and was the primary focus. Adding the MIR retrieval would add 7 additional years (1997-2004) of xCO2 data and also there is the potential for expanding geographically in future plans. E.g. Lauder has been measuring in this spectral range since mid 2020.

Since the MIR does not include O2 signatures pressure data have been used to calculate XCO2. Accordingly, the quality of the XCO2 product relies on the quality of the pressure sensor and its data. The paper would benefit from a paragraph or appendix describing the used pressure sensor and its requirements. In particular, the required precision and accuracy and any means to check and correct for drifts need to be discussed in the paper.
Great suggestion, a new appendix section will be added.

The paper includes an error estimate and the precision of the retrieved XCO2 product is good. Normally, a correlation coefficient of 0.95 (Fig. 11) is quite good, too. For XCO2 data however, it is not fully clear whether this is sufficient. For Burgos site, the correlation coefficient (0.84) is even lower.
Yes, the correlation is lower for Burgos, and we have two hypotheses as to why. Firstly, because the w4790 window has a higher temperature dependence, and secondly the position of the tropopause in the a priori profiles doesn't match the height of the real tropopause. These 2 factors combined result in a less accurate temperature profile leading to higher errors and lower accuracy of the retrieved xCO2. This accuracy is something that needs to be further improved.

It is certainly a good idea to include spectra from a wet site. The site used for this, however, suffers from the lack of NDACC spectra. Instead, spectra were recorded with a broadband InGaAs detector without any optical filter. In the 4800 cm-1 region an NDACC type spectrum recorded with a cooled InSb detector and bandwidth filter is better with respect to linearity and signal to noise ratio.
If this is the case then it will improve the accuracy of the retrieval in warm wet places. And it could be an additional explanation to the above discussed comment. Certainly it would be an interesting follow up study.

You might add the signal to noise ratio of the InSb and InGaAs spectra as shown in Figs. C1&C2. This might give a hint on the noise level although the spectral resolution differs strongly. Alternatively, did you record these spectra (C1&C2) with the same spectral

Resolution?
The signal to noise ratio has been added in the text . The spectra used for the smoothing is the same 3 years used in the rest of the paper, 2016 to 2018. So 6300 and w4790 InGaAs have the same resolution and they are retrieved from the same TCCON spectra. But the w4790 InSb has a different resolution (180 cm OPD) as it was measured with NDACC specifications.

The spectral resolution of 0.005 cm-1 as used in the NDACC mode is probably overdone in this spectral regime at 4800 cm-1. (At this spectral regime, my personnel guess would be something about 90 cm OPD or a resolution of about 0.01 cm-1, respectively.) The authors tested different spectral resolutions (Appendix D) which is a very useful exercise. However, I do not see a clear statement on the optimal resolution at 4800 cm-1. I would suggest adding a plot of a CO2 line recorded with different OPD to study the needed OPD to resolve the line fully.
I agree, this is why I performed the resolution tests. I added a plot with the suggested resolution. The final suggested resolution is the same as TCCON of 0.02 cm-1 because as you mention it is the resolution that allows the lines to be fully resolved.

In the introduction, the paper states that previously published recipes for XCO2 retrievals does not yield XCO2 data of sufficient quality to use the data for atmospheric research. This is correct. However, in the discussion or conclusion the results of the data set retrieved in this paper have not been compared to those previous data sets to demonstrate the improved data quality.
We have added a short comparison to the conclusion:

The retrieval proved to not have a big dependence on the a priori to correctly represent the daily and seasonal cycles (with correlation coefficients of at least 0.980 for all 3 retrievals:). **This is an improvement from Buschmann et al. (2016) where the averaging kernels limited the sensitivity making the retrieval highly dependent on the a priori. The retrieval also proved useful to retrieve daily and subdaily values, which is a different goal and purpose of Barthlott et al. (2015) study.** However…
Line 390

Furthermore, the long-term stability of a time series of this data product is not studied in detail. This should include a discussion of a possible drift of the pressure sensor. See comment above. Secondly, regular cell measurements to retrieve the ILS (Instrumental Line Shape) of the spectrometer are strongly recommended over the entire time series. Are ILS measurements made and used in this study and if so what are the ILS results? ILS parameters are also missing in the error estimate (Table 4).
The pressure sensor details and drift discussion has been added in Appendix F.
The ILS is not used in this study. The instrument is regularly aligned to meet TCCON requirements. This includes monthly cell measurements to monitor its performance. But indeed the ILS would affect the xCO2 retrieval, and that the magnitude of it should be characterised along with the ADCF when the full w4790 xCO2 product is produced.

While chapter 6 announces historical data from 1997 to 2018, Figs. 9, 10 and B1 just show data for the years 2016, 2017 and 2018. It seems this is limited to the availability of data from the 6300 cm-1 spectral region, isn't it? The NDACC type measurements at 4800 cm-1 started in 1997. When did the TCCON measurements at 6300 cm-1 started at Ny-Ålesund? It would be helpful for the reader if you add a table or columns to Table 1 listing the periods of available and used spectra for each retrieval and chapter. Also for Table 3 and Fig.7 the period covered would be helpful.

The historical data covering since 1997 is only in Figure 12.

The other plots in the time series, Figures 9 to 11, use the error weighted daily means for the periods 2016-2018 for Ny Alesund and 2017-2018 for Burgos as it is the overlapping years for the comparison.

In the rest of the paper's analysis, the xCO2 from InGaAs spectra (for both 6300 and w4790) use the same 3 and 2 years correspondingly.

But for InSb, only in section 4.2, I chose to use data from 1997 to 2018 to compensate for the lower quantity of measurements in comparison to TCCON. I have clarified the years of data used for each section. Also included in Table 1 is the actual availability of data for each spectra type.

There is a larger availability of data for 6300 as described in section 2.1 but it was decided to use a limited range for the study.

Finally, if earlier TCCON data (from before 2016) are available it would be nice to include a longer data set into the comparison. Moreover, to study the long-term stability in more detail it would be good to calculate the difference of NDACC and TCCON XCO2 data from the beginning of these measurements and to perform a trend study on the difference.

Initially, we tried to keep the time series consistent between Burgos (available only since 2017) and Ny-Ålesund and are confident that the number of data points allows for meaningful comparison. We certainly see the benefit in performing the analysis not only on longer time scales but with additional sites. However due to time constraints we suggest this effort be taken in a follow up study.

Technical corrections:

- l. 20: validation validation
- l. 20: ofCO2
- l. 25: sppectru, => spectrum
- l. 59: In NDACC spectra => NDACC spectra
- l. 126: Correct use of respectively?
- l. 141 and a few more times there is still a question mark
- l. 158: Ny-Ålesund around => Ny-Ålesund and around
- l. 163: following => following equation
- l. 169: when => even if?
- l. 185: => (Wunch et al., 2011a)
- l. 226: the those => those?
- l. 229: These bias => These biases or this bias
- l. 311: is and full stop are missing.
- l. 326: points. . => points.

- l. 407: Fig. D1 is missing or the number of Fig. D2 is incorrect.
- l. 412: where => were or better rephrase the sentence beginning with ‘In resolution Fig D1’
- l. 445: Ny-Ålesun => Ny-Ålesund
- l. 449: One authors => One author

All corrected.

The list of typos is not complete. Please proofread carefully.