General Comments

Reviewer's comments:

"So much of the abstract seems, well, abstract. Is this a validation paper, basically validating OCO-2 IAV so we can have more confidence in flux inversion results? Or to better understand the spatiotemporal scales at which we should aim our flux efforts which utilize OCO-2 data? I suggest making the abstract a bit more clear about how much the manuscript is validating OCO-2 IAV, versus to what degree it is doing interesting analysis with the IAV itself.

Regarding comparisons to TCCON & MBL sites: it seems like because of the low sampling associated with TCCON and the MBL sites, OCO-2 derived IAV is more powerful because of the better spatial sampling. You may wish to point this out in the abstract and/or conclusions more specifically. (I also wonder how much better future wide-swath sensors may be). Did you ever consider applying your method to GOSAT to derive IAV, to see how it compares to OCO-2? It would be especially interesting as we have 11+ years of GOSAT XCO_2*."*

Authors' Response:

We thank the reviewer for their generally positive review and for the constructive suggestions. We agree with the reviewer, there are many elements to the analysis we are presenting, and the paper consists of two parts: first, validating the usage of space-based OCO-2 detection by comparing the XCO_2 IAV based on the OCO-2 satellite against TCCON ground-based and MBL sites observation. Second, characterizing the spatiotemporal patterns of the interannual variation of the atmospheric CO_2 to understand its global drivers, since OCO-2 provides better spatial sampling than past datasets. In response to the reviewer's suggestion, we incorporate GOSAT observations available since 2009 in our analysis. GOSAT shows the similar pattern as OCO-2, and we are able to see that OCO-2 data show improvements in terms of the data quality, as expected.

Following the reviewers' suggestion, we add a bit more explanation for the purpose of the paper in the abstract/conclusion sections, and we describe the benefits of the improved spatial coverage of OCO-2. We also added figures and analysis based on the GOSAT dataset.

Key new lines in the abstract read "The similar zonal patterns of OCO-2 XCO_2 IAV timeseries compared to ground-based in situ observations and with column observations from the Total Carbon Column Observing Network (TCCON) and the Greenhouse Gases Observing Satellite (GOSAT) provide validation that OCO-2 observations can be used reliably to estimate IAV. Furthermore, the extensive spatial coverage of the OCO-2 satellite data leads to more robust IAV timeseries than those from other datasets, suggesting that OCO-2 provides new capabilities for revealing small IAV signals despite sources of noise and error that are inherent to remote sensing datasets. "

Specific Comments

Reviewer's comments:

Abstract: "The amplitude of IAV variations is up to 1.2 ppm over the continents and around 0.4 ppm over the open ocean." Please make it clear in the abstract that you are defining "amplitude of" as "standard deviation of".

Authors' Response:

We added explanation that amplitude is calculated as the standard deviation of the timeseries. The revised abstract reads "The IAV amplitude, calculated as the standard deviation of the IAV timeseries, is up to 1.2 ppm over the continents and around 0.4 ppm over the open ocean."

Reviewer's comments:

Sec 2.2.1: The results of the spatial aggregation sensitivity analysis seem to show substantial differences in the IAV depending on the spatial scale of aggregation. How do you know which spatial scale is most accurate, given that the differences are not just noise, but show large-scale biases? These large-scale differences clearly matter, as you show later most of the IAV is less than 0.75 ppm. I strongly suggest you repeat your sensitivity analysis with high-resolution model data (rather than real data), sampled like OCO-2. If you use model data, you know the right answer, so you can see what you can get away with. Something like the GMAO 0.75° model should have sufficient resolution for this purpose.

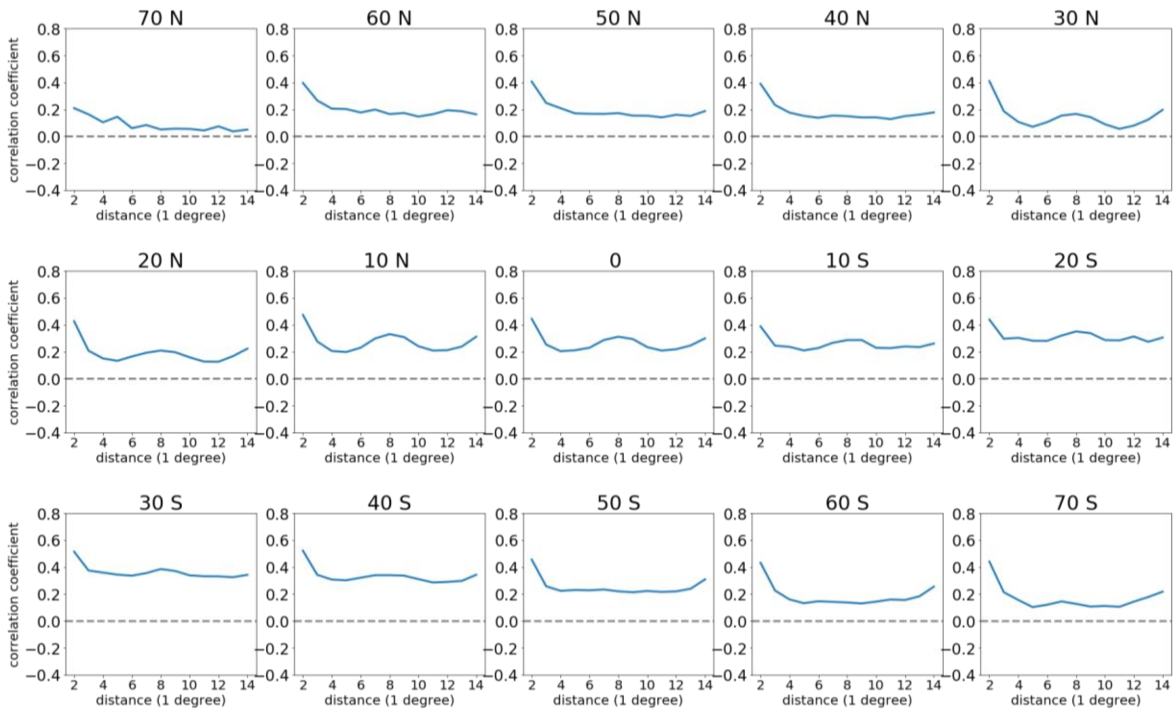
Authors' Response:

We agree with the reviewer that conducting a sensitivity analysis with high-resolution model data such as GMAO 0.75° would be interesting, yet we consider that the work involved in the suggested model analysis would first require a validation of the model IAV, and the scope of adding a model analysis is beyond the scale of our current study. Therefore, we prefer to keep this paper focused on observations only.

To address the reviewer's concerns, we conducted sensitivity studies with the observations themselves in which we aggregate the OCO-2 detection in different resolution - from 1° resolution to 15° resolution, and compare the difference of these resolutions (In Fig S1 and S2) to find the threshold that balanced the two goals of reducing noise yet revealing IAV at sub-zonal resolution. At 5°x5° (lower latitudes and 5°x10° (higher latitudes), the appearance of hotspots is minimized and the IAV amplitudes are spatially smooth, which we interpret as IAV signals emerging above the noise.

We further calculated the correlation coefficient among the IAV timeseries in neighboring gridcells of 1° resolution (Fig. S9). The R value decreases as the distance between the gridcell increases in each 10° zonal bands from 70°N to 70°S. We see a rapid decrease in the correlation between gridcells separated by 1° and gridcells separated by 5°, and generally stable correlation coefficient from 5° to 15°separation. Fig. S15 suggests that aggregations greater than 5° may over-smooth real variability.

We have included the new figure S9 in the supplement, and added the following text to the revised manuscript that reads "The 5°x5° and 5°x10° aggregation strike the necessary balance of reducing noise (evidenced by the smoother IAV amplitude fields as aggregation increases in Fig. S1) but maintaining spatial information by not oversmoothing (evidenced by the fact that the aggregation occurs at spatial scales finer than the "elbow" where correlations among 1° gridcells stop changing with separation distance in Fig. S9.



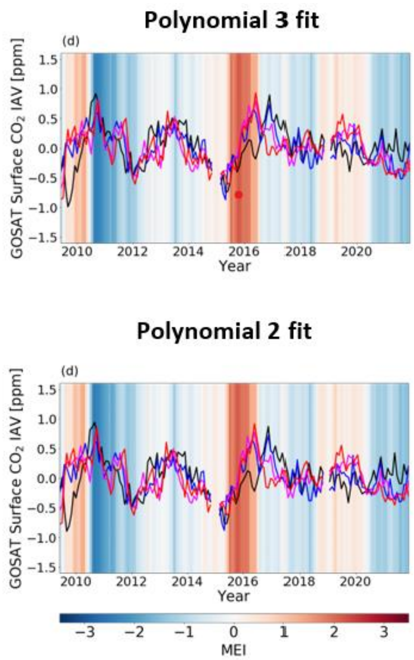
Supplementary Figure 8. Mean Correlation coefficient between the OCO-2 XCO_2 IAV of neighbouring gridcells in each 10° latitudinal band.

Reviewer's comments:

Sec 2.2.2: It seems like using a 3rd order polynomial on a 7-year time series to remove the secular increase is a recipe for problems when trying to derive the IAV. Wouldn't this artificially remove some of the IAV? Please discuss why 3rd order is necessary in the paper. Did you test 1st or 2nd order ? If so, why were they not sufficient?

Authors' Response:

We use a 3rd order polynomial to derive the long-term trend for OCO-2, TCCON, MBL sites, and newly added GOSAT in the revision. When comparing the 1st and 2nd compared to 3rd, we didn't see a quantitative difference between the IAV patterns for OCO-2 (newly added Fig. S9). Although 1st or 2nd is sufficient for 7-years OCO-2, GOSAT observation is from 2009 to the end of 2021, and half of the TCCON sites we use have data more than a decade. Therefore, to guarantee our methods closely captures their trends and confirm that we are applying the same methodology to all the datasets included in our paper, we use the 3rd order polynomial fit. We added text in the methods section to justify the 3rd order fit that reads "We fit a third order polynomial to the raw timeseries since the GOSAT,MBL and TCCON timeseries extend over a decade in length. We confirm that the use of a third-order polynomial, versus a second-order polynomial, does not remove the IAV signal from the shorter OCO-2 timeseries (Fig. S9)."



Supplementary Figure 9. Timeseries comparison between the zonal mean GOSAT XCO_2 IAV, based on detrending method using 2nd and 3rd polynomial fit.

Reviewer's comments:

Fig 5a: Care to comment on the strong feature near the beginning of 2020 peaking at 60°S latitude? That seems stronger that random variability.

Authors' Response:

In section 3.1, we mentioned that "For the two El Ñino periods in 2015-2017 and late-2018 to 2021, high IAV values originate in the tropics and a smooth transition to high IAV values is seen at higher latitudes as time progresses (Fig. 5a)." The strong feature at 60°S near the beginning of 2020 looks suspicious, since when we explore the NOAA ESRL sites - SYO and PSA - around 60°S, there is no such pattern. This peaking based on OCO-2 does not look like random variability, and no potential related geophysical events can serve as good explanations , for example, no documented volcanic eruptions or fire events that could have caused it. We explore if the anomaly peaking can persist after applying a more aggressive filtering - we additionally required that when calculating the average IAV of a certain zonal band of a certain month, there shall be observations for at least one third of the time - which is 24 out of 72 gridcells of available OCO-2 detection. With this aggressive filtering, the highly anomalous period was weakened yet still evident. We think the reasonable explanation of the peaking near the beginning of 2020 peaking at 60°S latitude is mainly due to OCO-2 issues with OCO-2 observation quality.

We have added the following text to the manuscript in Section 3.1, “While the OCO-2 patterns largely conform with variability expected based on ENSO and are in broad agreement with other observational networks, there are some anomalies that do not have an obvious explanation as of yet, such as the high XCO_2 in early 2020 around 60°S. Even with more aggressive data filtering, this episode persists, requiring more investigation of unknown geophysical drivers of high XCO_2 or potential retrieval issues that could cause a large positive bias.”

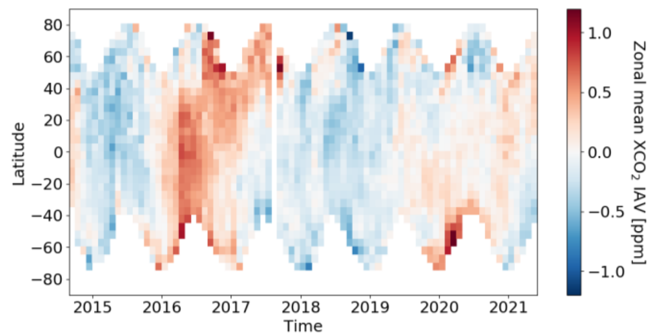
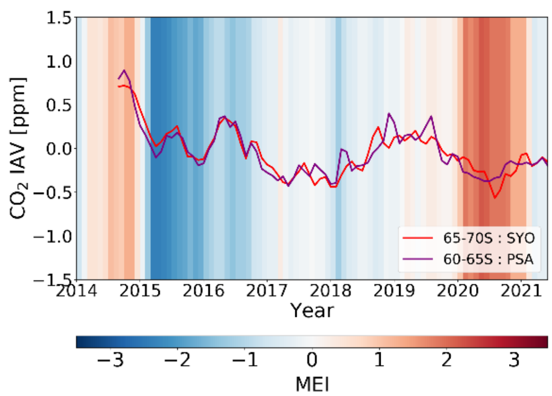


Figure 5. (a) Hovmöller Diagrams diagram showing zonal mean OCO-2 XCO_2 IAV timeseries for 5° latitude bins



IAV timeseries of NOAA ESRL sites around 60°S - SYO and PSA

Reviewer’s comments:

Near Fig6: Because MEI/ENSO is such a heavily discussed topic in this work, a plot of the correlation coefficient of MEI with IAV timeseries in local 5x5 grid boxes may be warranted – similar to figure 6. Have you made such a plot, and does it show any interesting teleconnections? You may need to introduce a lag at the more northern latitudes when calculating correlation coefficients there (a simple 0-6 month lag as a function of latitude could work).

Authors’ Response:

We added a map (new Fig. 8) showing the correlation coefficient between the IAV timeseries with MEI, corresponding to Figure 4 which demonstrates the relationship between zonal-mean IAV and ENSO. With no month lag, we were able to see the mainly positive correlation over both ocean and continents, both Northern Hemisphere and Southern Hemisphere. We added text to section 3.1 of the paper to discuss this figure. The new text reads “We assess the spatial correlation patterns between the IAV timeseries and MEI (Fig. 8a). The XCO_2 IAV timeseries have strong correlation coefficient with the MEI index in both Southern Hemisphere and Northern Hemisphere low latitudes from 0 to 30°N at lag 0, whereas in the Northern Hemisphere extratropic, the maximum positive correlation occurs at month 4 (Fig. 8b). The positive correlation between MEI and the IAV timeseries is gradually attenuated, with no clear correlation at six months lag (Fig. 8c). ".

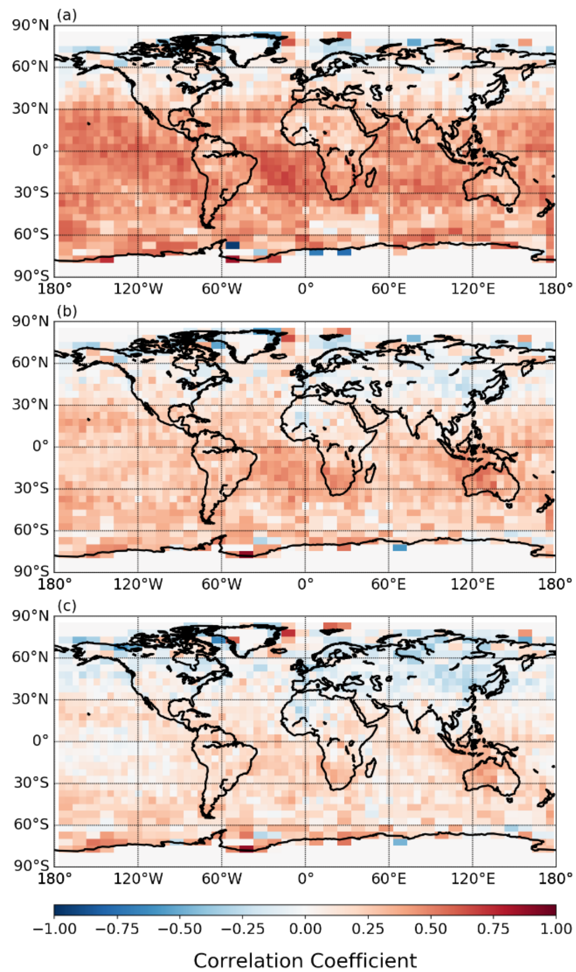


Figure 8. Correlation coefficient between local grid cell OCO-2 XCO_2 IAV timeseries and MEI .

Reviewer's comments:

Figure 10: Each "point" on the plot has in fact some uncertainty on the IAV at each site, due to both retrieval errors and spatiotemporal noise. Is it possible to get an estimate of this, and use it to add x & y error bars on each point? That might give a better picture of how consistent TCCON and OCO-2 IAV are, to within their respective errors. This figure implies that they are not very consistent. Related to the above, please check your IAV stddev calculations. I tried to reproduce your Bialystok numbers for OCO-2 and TCCON. Just by eyeballing your Fig S11, I got 1.05 for OCO-2 (similar to your number), but I got 0.75 for TCCON, whereas you got roughly 0.5. I wonder if some of your TCCON values are too low for some

reason (in particular at the NH sites). Your values at Karlsruhe and Orleans also both seem unreasonably low (both less than 0.4 ppm).

Authors' Response:

We add x & y error bar for TCCON & OCO-2 on each point for the IAV based on the year-to-year difference of the seasonal cycle, and we see sites in the Northern Hemisphere are affected more by the errors and noises. We checked the IAV standard deviation calculation and the plotting of Figure 10, and found the mismatch between the TCCON IAV amplitudes and their locations, after the correction of location, we could see the accordance between this Figure 10 and supplementary S11 which shows the IAV timeseries.

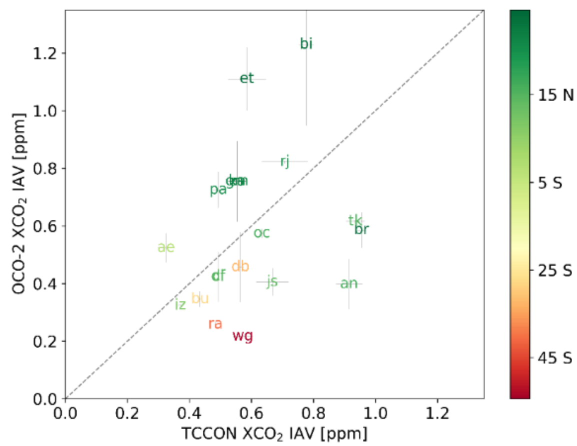


Figure 12. Comparison of OCO-2 and TCCON XCO_2 IAV amplitude at individual sites. Colours reflect site latitudes. The grey dashed line is the one-to-one identity line. The grey solid line is the error bar of the IAV amplitude.

Technical Comments

Reviewer's comments:

Line 72: Remove comma after "Chatterjee et al."

Authors' Response:

We correct the wrong format and remove the comma after *"Chatterjee et al."*

Reviewer's comments:

L76: "…is being used implicitly for flux attribution…". Please provide example references.

Authors' Response:

We cite the reference: Nassar, R., Jones, D. B. A., Kulawik, S. S., Worden, J. R., Bowman, K. W., Andres, R. J., Suntharalingam, P., Chen, J. M., Brenninkmeijer, C. A. M., Schuck, T. J., Conway, T. J., & Worthy, D. E. (2011). Inverse modeling of CO_2 sources and sinks using satellite observations of CO_2 from TES and

surface flask measurements, *Atmospheric Chemistry and Physic*s, 11, 6029–6047. https://doi.org/10.5194/acp-11-6029-2011.

Reviewer's comments:

L85: Please add reference Baker et al., Geosci Mod. Dev., https://doi.org/10.5194/gmd-15-649-2022.

Authors' Response:

We add the reference *Baker et al., Geosci Mod. Dev., https://doi.org/10.5194/gmd-15-649-2022, for "This is especially important in light of analysis which suggests that the error variance budget in OCO-2 observations is large and contains substantial spatially coherent signal".*

Reviewer's comments:

L110: Replace the O'Dell et al, 2012 reference with the O'Dell et al, 2018 reference. The former applies to GOSAT; the latter applies to OCO-2 and is a much more appropriate reference.

Authors' Response:

Thanks for pointing out misusage of citation, we replace the O'Dell et al, 2012 reference with the O'Dell et al, 2018 reference for the sentence "After filtering and bias correction, the OCO-2 XCO_2 retrievals agree well with TCCON in nadir, glint, and target observation modes, and generally have absolute median differences less than 0.4 ppm and Root Mean Square differences less than 1.5 ppm."

Reviewer's comments:

Sec 2.1.2: Please state somewhere if you use GGG2014 or GGG2020 (I'm assuming the former).

Authors' Response:

We specify the usage of GGG2014 with the sentence, "Data are publicly available from the TCCON GGG2014 Data Archive (https://tccondata.org/) hosted by the California Institute of Technology."

Reviewer's comments:

Sec 2.1.3: If there is any kind of version number or data source website for the NOAA sampling data, please provide it.

Authors' Response:

We add the data source of NOAA MBL sampling data: "To explore differences in surface and column-average CO_2 IAV, we analyze IAV in the surface CO_2 mole fraction at marine boundary layer (MBL) sites in the NOAA (National Oceanic and Atmospheric Administration) cooperative sampling network(https://gml.noaa.gov/dv/site/?program=ccgg)."

Reviewer's comments:

Figure 1: Most of the TCCON sites are in completely wrong places!! It looks like the longitudes are screwed up?

Authors' Response:

We express our thankfulness for the review pointing out the problem in mapping the TCCON locations.

We made corrections in the locations map Figure 1, confirming that TCCON sites in right places.

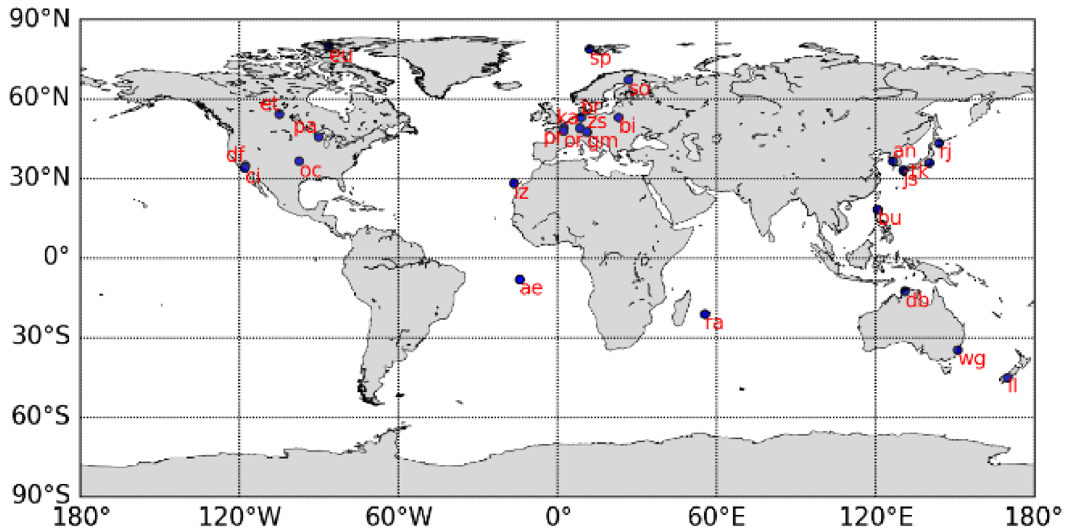


Figure 1. Map showing the locations and the acronyms of the TCCON sites.

Reviewer's comments:

L161: Please provide a reference for the NOAA monthly OLR data set, or remove the sentence about the source of the ENSO-related variables (which is probably not necessary as it will be given in the MEI documentation).

Authors' Response:

We remove the unnecessary sentence describing NOAA monthly OLR dataset.

Reviewer's comments:

Figure 9 caption: Do you show R or R2 in panel b? Please make the caption more clear. Currently the caption says R and the plot says R2 on the axis label.

Authors' Response:

We would like to show correlation coefficient R in Figure 9 panel b, the caption is now corrected.

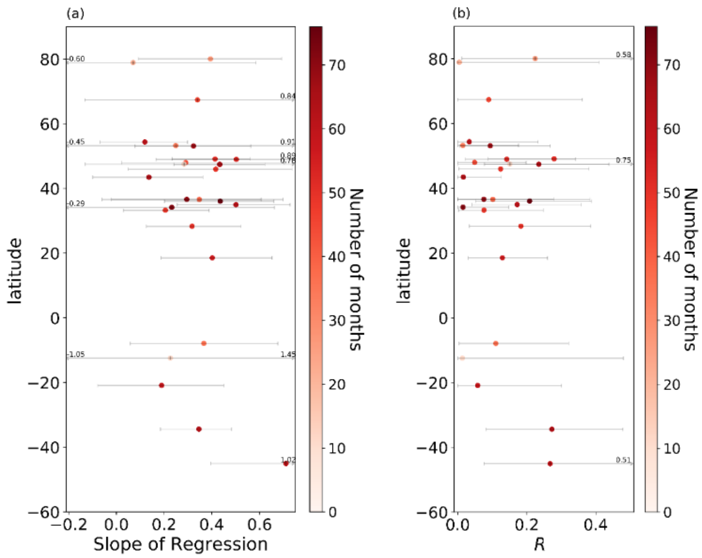


Figure 11. Latitudinal profile of regression Slope (panel a) and correlation coefficient (R, panel b) of OCO-2 versus TCCON XCO_2 IAV. The Slope and R values are based on using monthly XCO_2 IAV. The error bars result from a Monte Carlo bootstrapping approach . The colours represent the number of months data which used for the regression calculation, given gaps in both the OCO-2 and TCCON datasets.

Reviewer's comments:

Line 362: Change "R" to "correlation coefficients R". Otherwise, we really have to guess that you mean correlation coefficient.

Authors' Response:

We clarify the misleading sentence and now it reads like "We derive the regression slopes and Correlation Coefficient R between OCO-2 and monthly averaged TCCON IAV through bootstrapping Linear Regression fitting techniques to investigate the coherence between IAV signals from space-based and in-situ ground-based observations."

Reviewer's comments:

Line 387: "although we note that the IAV amplitude is a factor of almost two smaller in the column average mole fraction". Relative to what? Please add "relative to boundary layer CO_2" or something similar.

Authors' Response:

We clarify the misleading sentence and now it reads like "All datasets show consistent patterns in the response to the El Ñino periods, although we note that the IAV amplitude is a factor of almost two smaller in the column average mole fraction compared to the boundary layer CO_2, which reflects the fact that IAV variations emerge due to surface fluxes in the lower part of the atmosphere,…"

Reviewer's comments:

Fig S1 Caption: Please give the min # of soundings per gridbox.

Authors' Response:

We change the caption of Fig S1 into "The IAV amplitude map, with different resolution from (a) 2.5° longitude by 2.5 ° latitude , to (b) 5 ° longitude by 5 ° latitude, to (c) 10 ° longitude by 10 ° latitude , to (d) 5 ° longitude by 10 ° latitude and (e) 5 ° longitude by 15 ° latitude, each gridbox has at least 5 soundings."

Reviewer's comments:

Fig S3 Caption: Please give the spatial gridding (5x5, etc.) used, and change word "use" to "using" in the caption.

Authors' Response:

We change the caption of Fig S3 into "The IAV amplitude map, using different sounding numbers as the benchmark to filter and get the aggregated 5°x5° OCO-2 detected XCO_2."

Reviewer's comments:

Fig S6 Caption: Suggest changing word "record" to "years" ? It took me a while to figure out what you were getting at here.

Authors' Response:

We change the caption of Fig S6 into "The number of valid records (X out of 6 for JAN~JULY 5 for AUG or 7 for SEP~DEC) for each month (JAN, FEB, etc…) for each 5°x5° gridcell."