RESPONSE TO THE TCCON COMMENTS

We thank the TCCON community and the PIs for the valuable suggestions. Point-by-point responses are given below and marked by green color. For brevity, we respond "OK" to the comments that were precisely applied according to the recommendation and "reformulated" to the comments where we had to rethink about new formulation according to the less specific comment. Note that the definition of GR, AGR and MGR terms are given in the article.

– The paper does not seem to address the difference between the growth rate in the total column and the growth rate in the surface measurements. There is an important conceptual difference, which could be used to assess vertical mixing in the models.

Response 3.01: We have incorporated this analysis (see Figure 12 and the corresponding section).

– In general, it would be beneficial if the statistical results could be supported by some mechanistic analysis and facts about the ground-based locations. σ_a

Response 3.02: We have added the description of TCCON sites used in this study (second paragraph of Section 2.1.1). Analysis of selected TCCON sites is added (section 3.2.1).

– The method of determining the AGR could be more robust. The simple month(year+1) - month(year) method is prone to uncertainties introduced by short-term variability or (see below) non-temporally uniform sampling. At least using the median in calculating from the monthly growth rates should remove some of the extraneous variability, but the influence of irregular data on the robustness of the calculation of growth rate needs to be investigated. (See point a few below.)

Response 3.03:

- Since our paper is not aimed on the intercomparison of different methods for AGR, we strictly followed Buchwitz et al., (2018) procedure as the robust method. However, we have also acknowledged the raised issue from your comment as *"AGR may considerably vary depending on the methodology of calculation (Piao et al., 2020)"*.
- We did use median in calculating the monthly growth rates for every year. Basically, what the TCCON reviewer called "The simple month(year+1) - month(year) method" is MGR shown by Equation 3. Meanwhile, *"median in calculating from the monthly growth rate"* is basically AGR shown in Equation 4.
- The influence of irregular data is thoroughly investigated in this version of the article (see objective (a) of this study and the analysis for Figure 2, upper panel of Figure 3, Table 1, orange area of Figure 7)

– It would be good to include some details about selection criteria for the sites that are used. For example, from a TCCON perspective, why is Ny-Ålesund not included?

Response 3.04: Ny-Alesund is not completely excluded from the article (see Figure 5 and Figure 12 for instance). It is missing on Figure 6 due to specific reason *"ERK and NYA stations are not used in this analysis due to lack of data for evaluating seasonal cycle according to LM."* where LM is Lindqvist et al., (2015) methodology. All details for filtering out TCCON stations are given on Figure 2 and section 3.1.

– It would be beneficial to include some kind of sampling bias treatment in your analysis to account for the sometimes-sparse time series. Some further comments on this follow.

– The paper presents a highly-averaged statistical inter-comparison of TCCON data-inferred atmospheric growth ratio (AGRTCCON) and its variability with models and satellite methods. One key finding is that the AGRTCCON correlates with models and satellites only after 2010. This is attributed to the expansion of TCCON after 2009, agreeing well with the references only during the 2010s. This hypothesis should be tested by looking at some specific long-term sites, recognizing that effects of the seasonal cycle are important to disentangle from the broader analysis. This can be reduced by picking a southern or low-latitude TCCON, and selecting for a restricted range of solar zenith angles. It would be worth the authors considering the analysis by Lindqvist et al. (2015). In their analysis, they fit TCCON data using a time-varying 6-parameter function for the northern hemisphere. Presenting such a fitting analysis would strengthen the mechanistic interpretation of the paper, for example helping to determine which TCCON sites are best suited to capture variability in the AGRs dominant global effects like ENSO.

Response 3.05.

- We analyzed the extended time series including 2018 and the hypothesis about the period-driven agreement is rejected in this version. See the new explanation of the TCCON-to-reference agreement part (Section 3.3).
- Please note that we have invited Hannakaisa Lindqvist to our manuscript. She helped us to make a comparison between the AGR and the linear trend of XCO_2 using their methodology (Lindqvist et al., 2015 paper). See Figure 6 please. It did help providing mechanistic explanation of which sites are more independent from the effects of time series edge/seasonal maximum CO_2 effect.
- The key sampling bias is the lack of station-wise AGR in the global AGR estimates. This irregularity is reflected at the error bars plotted using data abundance-driven uncertainties using *"average standard deviation across TCCON stations (AGR) multiplied on* $\sqrt{N_{total}/N}$ *factor where N - number of*

stations used, N_{total} - total number of stations in TCCON analysis." This calculation goes in line with Buchwitz et al., (2018) methodology.

– Another important finding is agreement for the large amplitude of the AGR in 2015-2016, during the strongest ENSO. This likely results from the high signal relative to the background in AGR across TCCON. The key point that should be clarified is not only the increase in number of sites, but also the locations of the sites and the timing matters in consistently deriving AGRs. Some fitting of individual TCCON datasets could strengthen this statistical inter-comparison.

Response 3.06: Please see new Figure 6 for the fitting analysis of individual TCCON sites. We included the phrase *"In future, not only the increase in number of sites but also the locations of the TCCON sites and the timing of observations would be crucial for deriving robust CO_2 growth rates."* in the conclusions.

– Many TCCON sites are influenced by local/regional effects that would allow differences to be explained, and inform sub-sampling strategies to strengthen your comparisons. For example, at Manaus (Brazil) under the dominant influence of local rainforest, large daily drawdowns (1.8 ppm) occur, which could cause local sink signals on the order of the AGR.

Response 3.07: Manaus station is not included in the analysis due to lack of MGR for the period of study (see Figure 2 and the corresponding section).

Four Corners TCCON site was located near a power plant, which led to early morning plumes that increased xCO2 by up to 10 ppm (see and cite Lindenmaier et al. (2014), which could lead to biases in the derived AGR. These effects should be discussed, and e.g. for Four Corners only afternoon data used.

Response 3.08:

- Four Corners has too short time series to be included in the analysis (see the response 3.07 above for the details).
- We have included the investigation of the role of neighboring urban sources in the manuscript using MODIS urban pixels (see section 3.4.2 please). However, this method does not allow tackling strong non-urban sources such as power plants since standalone power plant facilities are likely missing in the MODIS urban data. Due to this, we added we included the warning about this in the conclusions *"strong CO_2 sources are not always attributed to built-up zones with clear spatial extent and structure such as megacities. They can be attributed to emissions from a power plant as it was shown in the previous studies (Lindenmaier et al., 2014)".*

We suggest having a short discussion on the various TCCON site locations, and local/regional effects, which would allow you to explain differences, define sub-sampling strategies to avoid local/regional effects that bias comparisons, and provide more robust comparisons.

Response 3.09:

- We have included discussion of XCO_2, MGR and AGR series for several selected sites in section 3.2.1 (Figure 4).
- Since the role of data abundance/sampling/irregularity for AGR calculation is pivotal, we have added a separate objective to investigate this issue. Namely, (a) *"to estimate the robustness of AGR_{TCCON} due to the data sampling, measurement gaps or difference in time series across the sites".*
- The (a) objective is now supported by extended evidences about daily, monthly, annual and seasonal stability of TCCON AGR estimates (see 2nd paragraph of the discussion for the details). Since we found low influence of daily, monthly and annual data abundance (and seasonality) on MGR and AGR estimates, we did not apply more sophisticated data sampling strategy and concluded that *"current estimates of CO_2 annual growth obtained from the TCCON aggregated observations is adequate and are in reasonable agreement with the existing references even when the simple methodology (Thoning et al., 1989) and the simple TCCON data screening are applied"* (see conclusions).

You do discuss some of the model-TCCON site difference in MGR, specifically for Tsukuba, Ascension Island, and Pasadena. However, this discussion and the associated hypotheses are not well backed up by evidence. It also needs to be clear that the failure to represent the MGRs is presumably due to failure to capture something that varies interannually. Regarding the hypotheses, e.g. why would such a small landmass as Ascension Island have an impact at the model resolution, especially in total column space? Given its remoteness from large landmass, it would perhaps be expected that Ascension should be among the best represented sites, if we assume that the models get the land source/sinks slightly wrong. Do the disagreement correlate with particular modules within the models? We do know that Ascension can sometimes be influenced by biomass burning from Africa and/or South America. Perhaps this is a more likely driver of the disagreement. Or it could be that the models don't capture the interannual variability in terrestrial carbon exchange in either Africa or South/Central America - there are few other sites that could be influenced by this. For Pasadena, the conclusion that carbon uptake at high-latitudes could drive the MGR differences is wholly unconvincing, as there are many sites where the influence of such an effect would be much larger.

Response 3.10:

- Note that Pasadena and Tsukuba early hypotheses are now rejected based on the new evidences.
- We also think that MGR can exhibit disagreement due to exposure of TCCON site to some localized influence which cannot be captured by models.
- This hypothesis was checked by analyzing the urban influence on interannual TCCON signal is expressed by the agreement rate of MGR estimates between TCCON and the models (section 3.4.2). TCCON-to-model correlation coefficient has negative agreement (r = - 0.73) with the size of the closest megacity to TCCON station (calculated by MODIS urban pixels, megacity = city > 1500 km^2). Since it can indicate a potential exposure of TCCON station to urban CO_2 emissions, we approximately quantified this exposure by using the distance to the closest megacity. According to this analysis, Paris, Tsukuba, Saga, Pasadena and Karlsruhe are theoretically the most influenced stations (< 40 km to megacity). However, the reviewer concern about malignant role of these sites in global AGR signal is not major factor here since "The difference between original AGR_{TCCON} and AGR_{TCCON} without "the most urbanized sites" ranged from negligibly low ~0.00 ppm to 0.29 ppm (2017) despite these "most urbanized sites" composed >20% of observational cover of TCCON in 2017."
- Thank you for valuable comment about Ascension as this hypothesis about biomass burning is likely the reason of disagreement. Namely, we included the analysis of biomass burning fluxes using CT (see Figure 14 and the associated description in the paragraph above). Also, detailed time series of XCO_2, MGR and AGR from Ascension is added to supplement (Figure S.2.7).

At Tsukuba, you don't appear to make any real conclusion, but leave the peak carbon uptake hanging as a possible driver. If that were the case, surely neighboring sites (Rikubetsu, Saga) would exhibit similar disagreement. The effect might be related more to failing to capture topographical features around the site in the relatively coarse resolution models. There is perhaps a role here for using the satellites to break up the surrounding area to look at potential spatial effects.

Response 3.11: As mentioned in the response 3.10, we have not found evidences for the role of Tsukuba, Saga and Rikubetsu. The explanation of this disagreement is given in the same response 3.10. There is no high correlation between TCCON-to-model correlation coefficient (for MGR) and the altitude of the site. Despite this and the finding about urban influence, we incorporated the TCCON reviewer's suggestion as *"It should be noted that except the urban influence, the disagreement at some TCCON sites can be related to failing to capture topographical features around the site in the relatively coarse resolution models."*

– Please include in the introduction as well as late in the TCCON section a brief discussion of the outcome of the work of Yuan et al. (2019), which looks at comparison of in situ, TCCON, satellite measurements and CarbonTracker for (X)CO2

Response 3.12: Yuan et al. (2019) description is included in the introduction (3rd paragraph) with the following reference:

Yuan, Y., Sussmann, R., Rettinger, M., Ries, L., Petermeier, H., and Menzel, A.Comparison of Continuous In-Situ CO_2 Measurements with Co-Located Column-Averaged XCO_2 TCCON/Satellite Observations and CarbonTracker Model Over the Zugspitze Region, Remote Sens. **2019**, 11, 2981; doi:10.3390/rs11242981

– As indicated in our other comment, please ensure the TCCON data DOIs are appropriately cited for the sites used in this work.

Response 3.13: TCCON data DOIs are cited for all sites used in this work.

– Uncertainties? There are attempts to interpret sometimes small differences, but no attempt to quantify the uncertainties in the AGR or MGR estimates. In many cases, reporting to 2d.p. (3s.f.) might be excessive.

Response 3.14:

- We estimated the error spread of MGR input into the AGR (Figures 3 and 7). As only the threshold of "2" measurements is used in the manuscript, this estimate cannot be included in the error propagation but serves a good illustration for potential error spread could be caused by this choice.
- Regarding uncertainties stemming from the input data for AGR calculation, we chose a strategy similar to Buchwitz et al., (2018) approach for calculating their standard deviations for global AGR. Namely, as shown in Figure 8 *"Error bars of AGRTCCON (orange vertical dashed line) are defined based on the station-wise variability using method similar to Buchwitz et al., (2018). Namely, average standard deviation across TCCON stations (AGR) multiplied on* $\sqrt{N_{total}}/N$ *factor where N - number of stations used,* N_{total} *– total number of stations in TCCON analysis."* Due to this, on Figure 8 uncertainties of the years with lower number of TCCON observations from lower number of stations are higher.

– Overall, it seems that there are times in the paper where it is a case of the cart driving the horse. Of course it is important and valuable to understand the growth rates and their differences between different datasets, but really what we want to do is use these to diagnose what in our understanding is incomplete and leads to the differences. That's presumably some combination of surface fluxes and atmospheric processes. It is a subtlety, but at times the emphasis throughout the paper of quantifying growth rates is overstated, and should be rephrased to state their importance in interpreting the underlying biogeochemistry/physics.

Response 3.15: Perhaps, it was also a problem of a slight mismatch between the research aim formulated stated and the contents we present in the first version of the

paper. Please see that we have reformulated the research aim and the objectives as shown below.

- New research aim is “Our study aims to assess the robustness of GR estimates from the observations of Total Carbon Column Observing Network (TCCON) considering the existing spatio-temporal gaps of the network”.
- New objectives are “(a) to estimate the robustness of AGR_{TCCON} due to the data sampling, measurement gaps or difference in time series across the sites. Secondly, (b) to examine the AGR_{TCCON} agreement with the existing CO_2 growth references and its sensitivity to external factors Thirdly, an additional objective is set (c) to assess the exposure of MGR_{TCCON} estimates at each TCCON station to external factors”.

Despite it is critical to diagnose missing processes behind CO_2 growth, the main aim of this article is driven by not less important motivation. Namely, as it is stated now in the introduction *“The CO_2 growth rate (GR) is a relatively well-known quantity but there are few global observational approaches suitable for quantifying global GR. Our study aims to assess the robustness of GR estimates from the observations of Total Carbon Column Observing Network (TCCON) considering the existing spatio-temporal gaps of the network”*

Minor Comments/Technical Corrections

The article requires careful proof-reading as there are many grammatical errors. We have attempted to report these, but the list is probably not comprehensive.

Response 3.16: However, this version of the manuscript has undergone considerable revision and the language was a subject of the revision as well. We incorporated the language corrections according to the recommendations from the English native speakers from the TCCON community side. For instance, language-related suggestions of Nicholas Deutscher provided in the collective comment and the language corrections by Debra Wunch in the personal email correspondence.

– line 9: remove "the" before "global warming" - OK

–line10-11: "Despite atmospheric CO2 growth rate had been considered as the well-known quantity" This wording doesn’t work. Maybe "Despite the atmospheric CO2 growth rate having been considered as a well-known quantity" - Reformulated

– line 13: Please correctly define TCCON as the "Total Carbon Column Observing Network" - OK

– line 23-25: The structure of this sentence is confusing, and forms a double negative. It would perhaps be better to emphasize good agreement at 85% of stations. - Reformulated

– line 27: "perfect (r=0.99)" - not quite perfect – r = 0.99 correlation is denoted as "excellent"

– line 29: in-> a; also again, this is not "perfect". OK

– line 29: insert "a" before "spatial" - Reformulated

– line 33-36: sentence seems back to front. Missing 'a' before CO2 - Reformulated

– line 42: "permanent" - we'd like to hope it isn't permanent, and while perhaps that is optimistic, a better word should be chosen here – "Persistent" instead of "permanent" is probably better choice?

– line 43-44, and other locations: when referring to growth rate (GR) it should almost have "the" before it. E.g. here, it should read "The atmospheric CO2 growth rate (GR)". There are many instances of this throughout the manuscript that we will not explicitly point out each time. – Corrected to "The atmospheric CO2 growth rate" when this term is used in the text.

– line 46: "the precision of direct observations are high (0.09 ppm)" Please add what kind of instrument has such high precision with citations. In situ measurement? Picarro? Also, "high" is ambiguous, and would perhaps be better replaced by something unambiguous ("excellent") or re-wording ("direct observations are highly precise").

Response 3.17. We reformulated it in this way: *"As the direct observations of CO_2 by infrared analyzers in the background atmospheric conditions are precise (accuracy > 0.20 ppm), the GR in the entire atmosphere is known with high confidence."* We used (Dlugokencky and Tans, 2018) as a reference and information about 0.20 ppm precision is taken from the latest respective website (https://www.esrl.noaa.gov/gmd/ccgg/about/co2_measurements.html).

– line 47: "uploaded to" - a better phrasing might be "reported in" OK

– line 48: at-> "on a" OK

– line 50: than-> that OK

– line 51: "constrains" can have a specific meaning in flux estimation, perhaps replace with "limits" OK

– line 61: insert "the" before terrestrial OK

– line 67: process-> Processes Reformulated

– line 70: insert comma before which, remove "per se" OK

– line 71: ecosystem-> Ecosystems OK

– line 72: semiarid-> semi-arid - OK

– line 73: "Despite the regions"-> "Despite the fact that the regions" Reformulated

– line 74: evidences-> evidence (and are ->is) Reformulated

line 79: "the limited"-> "a limited" an alternative is to qualify "the limited number of stations at which measurements are available on long time scales" or something similar - Reformulated

– line 79-81: do we really want to know the growth rate everywhere? You probably want a growth rate for each distinct region (ecosystem or whatever). Your desire to know it everywhere perhaps points to a limitation of the study.

Response 3.18: We are interested in the global growth based on aggregated TCCON measurements. Please note that we reformulated the main research aim and objectives of the study accordingly (below). This change has been motivated by multiple reviewers' comments about research aim-results inconsistencies.

- New research aim is "Our study aims to assess the robustness of GR estimates from the observations of Total Carbon Column Observing Network (TCCON) considering the existing spatio-temporal gaps of the network".
- New objectives are "(a) to estimate the robustness of AGR_{TCCON} due to the data sampling, measurement gaps or difference in time series across the sites. Secondly, (b) to examine the AGR_{TCCON} agreement with the existing CO_2 growth references and its sensitivity to external factors Thirdly, an additional objective is set (c) to assess the exposure of MGR_{TCCON} estimates at each TCCON station to external factors".

Each site's growth rate can be affected by local factors, that may or may not be of interest for the growth rate, and not on a global scale. By selecting a simple 12-month difference method, you leave yourself vulnerable to transient effects from local signals, as does failing to sub-sample measurements to remove local effects. Of course there is a role for both understanding the regional scale effects and changes in local factors as well, but you need to be clear about what you are trying to achieve here.

Response 3.19:

- The example of the role of local factors influencing global growth rate is shown in section 3.4.2. The promising finding about low sensitivity of global

CO_2 growth to such strong local influence as exposure to large city CO_2 emissions indirectly points to stability of the current methodology and the use of simple sub-sampling strategies.

- The influence of data irregularity is thoroughly investigated in this version of the article (see objective (a) of this study and the analysis for Figure 2, upper panel of Figure 3, Table 1, orange area of Figure 7).

– line 85: ratio-> ratios - OK

– line 88: remove semicolon - OK

– line 96: add "the" before "Global Carbon Budget" – Corrected throughout the text

– line 100: "the TCCON"-> "TCCON's" Corrected throughout the text

– line 101: "approve"->"prove" Reformulated

– line 102: "rest"->"remaining" Reformulated

– line 105: "At second"-> "Secondly" (similarly next line for "At third" Corrected throughout the text

– line 108: "Sect.s"->"Sects" OK

– line 113: Spell out Section here, as you are not explicitly referencing a separate section. OK

– line 118: the measurements are not continuous because they rely on the presence of sunlight, so by nature cannot measure at night or in cloudy conditions. Quasi-continuous maybe, or ongoing OK

– line 119: "The XCO2 estimates are retrieved using the ratio" XCO2 is not retrieved. Replace with "The XCO2 values are obtained by taking the ratio" OK

– line 120: O2 (subscript) OK

Response 3.20: Most corrections from the fragment above applied, we did not apply the corrections only to those formulations disappeared or reshaped as a result of the revision.

– line 121-123: You need references here for these statements.

Response 3.21: The reference for several statements is Wunch et al., 2011 work. We mentioned this reference only once in the end of the entire block that belongs to this reference. More specifically, the statement about insensitivity of column-averaged dry-air mole fractions is given in the introduction of Wunch et al., 2011. Lines 11-12 of their work: *"Column-averaged dry-air mole fractions (DMFs; denoted XG for gas G) are particularly useful for this purpose because they are insensitive to variations in surface pressure and atmospheric water vapour".*

– line 123: "spectra by pointing on the sun in the near-infrared spectrum" "spectra in the near-infrared region by pointing at the sun" OK

– line 123: superscript OK

– line 125: delete "XCO2"; accuracy-> uncertainty (2-sigma) OK

– line 126: "calibrated" is actually the wrong word here, because it is not truly a calibration. They are compared to these measurements, which links them to the WMO scale OK

–line 127: "from World Meteorological Organization onboard" "traceable to the World Meteorological Organization scale" OK

– line 130: exist->existing OK

– line 131: "familiarize"->"familiarize themselves" OK

– line 133: "addresses to"->uses OK

– line 135-137: not clear what this sentence means - needs clarification -Reformulated

– line 145/Figure 1: In this figure, it appears that the green dot "ZUG" should in fact be "KAR" for Karlsruhe. ZUG and GAR are practically co-located and would therefore not be differentiated on this scale map.

Response 3.22: Figure is remade

– line 150: Insert "The" at the start of the sentence, and "is" before "being" OK

– line 151: "developed by the Institute of Pierre Simon Laplace (LSCE)" LSCE stands for "Laboratoire des Sciences du Climat et de l'Environnement" OK

– line 157: Sentence starting here needs clarification/rewording. Reformulated

– line 160: insert "the" before biomass OK

– line 161: insert "the" before "fire module" and needs clarification on GFAS (GFAS emissions possibly, or "uses the GFAS") OK

– line 164-165: how did you judge that "this process did not cause serious error propagation"?

Response 3.26: We omitted the comparison figure from the supplement due to similarity of the models' resolution but can insert it back if needed. This fact is also seen by the similar agreement between TCCON-CT and between TCCON-CAMS where the agreement with CAMS is not deteriorated. Also, by high spatial agreement between CAMS and CT on Figure 13.

– line 175: cover->coverage OK

– line 177: "CTA" What does this stand for? Replace with "CT"? OK

– line 177-179: please be clear that there are multiple modules for each flux within CarbonTracker. E.g. while CASA is used in both, there are two versions of the biosphere model.

Response 3.27: CASA includes GFED 4.1s (hourly resolution) and GFED_CMS (daily resolution)

– line 178: "GFED" is already defined in line 161 OK

– line 188, 221, 224: subscript $_2$ OK

– line 191: "being updated at"->"is updated on" OK

– line 192: why "Obviously"? Suggest removing this word. OK

– line 195: on-> at; insert "The" before GCB OK

– line 196: The correct journal name is Earth System Science Data OK

– line 197-199: Sentence needs to be revised. Either "Large scale AGR estimates... are also taken from" or "Another reference ... is the satellite data" (i.e. delete "are taken from" in this second option) OK

– line 201: Should be "Greenhouse gases Observing SATellite" OK

– line 210: is there also a neutral classification (0)?

Response 3.28: True, the sentence is corrected.

– line 226, line 577: Piao et al., 2019 –> Piao et al., 2020 OK

– line 240: about the definition of AGR, MGR was a difference between 2 successive years. But for AGR, isn't it also a difference between 2 successive years? Clarify the difference

Response 3.29: Please see updated more detailed description of MGR calculation (Equation 3) and AGR calculation (Equation 4).

– line 252: How do you determine the number of observations per month. Some sites routinely make several hundred per day, so presumably some pre-processing is done.

Response 3.30:

- The number of monthly observations is defined based on the availability of mean daily XCO_2 estimates from a TCCON station (the details about this procedure are shown in Figure 2 and section 3.1).
- Thanks for pointing out the irregularities in daily data abundance across the TCCON sites. Based on this, the current version of the paper analyzes the role of summarized sub-daily observations performed at each TCCON sites. For details please see Table 2 and the associated description in the paragraph above. The total number of daily observations is for each station is also illustrated on Figure S.2.2 (supplementary).

– line 252: less -> fewer Also, this still seems fairly loose, would it not be better to raise the minimum number of observations contributing to the monthly median? This is discussed later, but there is no reference to that later discussion.

Response 3.31: Raising the minimum daily threshold leads to wiping out many TCCON stations as shown in the section 3.1 and Figure 2. At the same time, the difference between the minimum amount of the sub-monthly input and somehow medium amount is surprisingly low (error spread from Figure 3 and orange area from Figure 7). Due to this, increasing the daily threshold does not lead to improvement between AGR_{TCCON} and the references as we made a note about this in the study: *"change of daily threshold from '2' to higher values (3-25) does not lead to improved correlation of AGRTCCON with either SAT ($r \leq 0.60$ for any threshold) or GCB ($r \leq 0.56$ for any threshold). At one hand, these findings confirm the correctness of our approach of saving as much data as possible by using daily threshold '2'."* In other words, we have not found any evidences that raising the number of the observations in the minimum threshold would strengthen the analysis and use the '2' threshold.

– line 266: "from 1.71 ppm (2009) to 3.35 ppm (2008)" => "from 1.71 ppm (2010 or beginning of 2011?) to 3.35 ppm (2009 or beginning of 2010?)"

Response 3.32: This sentence is reformulated according to the new results. We checked the consistency of AGR estimates with the years mentioned in the text.

– line 272: do you try selecting the model or satellite data to match the TCCON spatio-temporal measurement pattern? This would confirm or refute your assumption.

Response 3.33: We use satellite global growth estimates from previous study as the global-scale reference (Buchwitz et al., 2018). Another reference is from the global carbon budget (Friedlingstein et al., 2019). Moreover, the hypothesis about the period-driven agreement between AGR_{TCCON} and the references is rejected based on the results from the current version.

– line 273 "a study period"->"the study period" OK

– line 275: r = 0.61 for TCCON-SAT, r = 0.49 for TCCON-GCB In the Abstract and Discussion sections, the correlation coefficients are described as being 0.75 for TCCON-SAT and 0.68 for TCCON-GCB. Which is correct? Reformulated based on new results

– line 278: "right below"->"directly below the" OK

– line 281: not sure why you would expect vegetation-driven seasonality to result in wave-shaped fluctuations. Interannual variability, perhaps, but by the way you are calculating it you are accounting for the seasonal cycles.

Response 3.34: That is true. Seasonal-dependent analysis of MGRs at the selected sites confirms the lack of vegetation-driven seasonality (section 3.2.1). *"We report the estimates of seasonal MGRs in the brackets using the following order: winter, spring, summer and autumn (TSU = 2.57, 2.02, 2.54, 2.70 ppm; GAR = 2.62, 2.19, 2.26, 2.21 ppm; PKF = 2.03, 1.80, 2.24, 2.56 ppm)".*

– line 283: approve? Not the correct word here Reformulated

– line 287: "(right axis of Fig. 1, panel b)"->"(right axis of Fig. 2, panel b)" OK

– line 291: "to data"-> "with data" Reformulated

– line 292: remove comma Reformulated

– line 295: "Besides 15-year pattern" - not sure what you mean? Reformulated

– line 298: "yield to"->"yield a" OK

– line 300: This underscores the importance of comparing apples with apples, basically. That is, if you want to make a comparison between similarities in behaviour either across time periods or between datasets, then they need to be sampled to minimize potential spatio-temporal biases.

Response 3.35:

- To ensure there is no pitfall with "apples" here, we analyzed the role of daily, monthly and annual data to station-wise and global AGR_{TCCON} estimation (Table 1 and the corresponding description above).
- The uncertainty in Figure 8 (what was Figure 2 in the previous version) reflects the data input irregularity driven by varying number of stations used for calculation global AGR_{TCCON} for every year. This is a similar step to Buchwitz et al., (2018) approach of AGR uncertainty calculation where instead of months we use number of stations. As methodology subsection and Figure 8 caption state for instance *"average standard deviation across TCCON stations (AGR) multiplied on* $\sqrt{N_{total}}/N$ *factor where N - number of stations used,* N_{total} *- total number of stations in TCCON analysis"*

- We also followed the TCCON reviewer's recommendation from the comment with response 3.38 (below) and analyzed relationship between abundance of MGR for every year and TCCON-to-reference bias (see the response 3.38).
- To be aware which of our stations are more affected by seasonality or time series edge effects, we compared AGR with the six-parameter-based linear trend calculated by Lindqvist et al., 2015 methodology.

– Figure 2 and its caption: the scale used for the lines corresponding the AGR_XCO2 is not exactly every 8th month (August) as mentioned in the caption, for example for the peak of the gray line (AGR_TCCON) which is located around 2nd month of 2010.

– Figure 2 caption - missing subscripts on the AGR terms.

Response 3.36: This figure is corrected

– line 312: "during the year where periods starting from winter months are shown with green tones, from spring with gray tones, from summer with golden tones, from autumn with red tones." Clarify about differences for different hemispheres. For once, this seems to have a southern hemisphere bias->

Response 3.37: This does not seem to be southern hemisphere bias according to new Figure S.2.3 (supplementary material) where we split all MGR to northern and southern hemispheres according to the recommendation above.

– line 319: "be risen"->"arise" OK

– line 320 "(based on one-time TCCON observations)" Is it one measurement of TCCON or is it a daily mean measurement of TCCON? Reformulated

– line 321-322: Not sure what this sentence means - clarify. Reformulated

– line 329: "different thresholds we mentioned above"->"the different thresholds mentioned above" OK

– line 332-335: perhaps it would be better to correlate the number of available MGR with the ratio of TCCON to SAT (or model) error spreads, or again compare by subsampling the model or satellite data to resemble the relevant TCCON data for each period.

Response 3.38: We followed the recommendation and correlated the number of available MGRs (and AGRs) with the ratio of TCCON-to-SAT and TCCON-to-GCB bias for all possible daily thresholds. There is no strong correlation for MGR_{count} vs $TCCON\text{-}to\text{-}GCB_{bias}$ comparison ($r = 0.22 – 0.41$ depending on daily threshold) and for MGR_{count} vs $TCCON\text{-}to\text{-}SAT_{bias}$ comparison ($r = 0.25 – 0.46$). For daily threshold

of ‘2’ the correlation coefficients for the aforementioned comparisons are 0.34 (vs TCCON-to-GCB) and 0.44 (to-SAT) respectively.

– line 347: such as 30 points Reformulated

– line 349: approves->proves OK

– line 363: versus->with OK

– line 367: add "the" before "correlation" Reformulated

– line 378: plain->simplistic (assuming this is what you mean) OK

– line 381: regarding the Tsukuba site - probably the topography cannot be resolved at these model resolutions. Also, if the difference between models and TCCON is driven by the land carbon sink as you hypothesize, surely the neighboring sites (Rikubetsu, Saga) should exhibit similar behavior.

Response 3.39: Please see the 3.10 and 3.11 responses as the comments behind these responses have almost identical suggestions.

– line 387: It’s not "most likely" that Ascension is a small island-> Please fix this wording. Though as noted in the general comments it is not clear that this small landmass would have affect model-measurement differences here in the regionally-representative column.

Response 3.40: We fixed this awkward wording. Please see the additional analysis of biomass burning fluxes role incorporated at Figure 14 and the corresponding section above. Ascension is analyzed in details in this version of the paper (section 3.4.3 plus supplementary Figure S.2.6).

– line 397: add "themself" after familiarize OK

– line 400/Figure 4: It looks like CAMS is missing for WOL, and CT missing for ASC. It would be helpful to include the values for the coefficients on the plot.

Response 3.41: Please check new version (Figure 12) with correlation coefficients plotted at the edge of the bars.

– line 415: on-> by? OK

Seems fundamentally like the average AGRs should be in good agreement, unless there is something wrong with a dataset, or the comparative data are sampled so as to introduce biases between them. This is again an instance where the models/satellites should be sampled in the same spatio-temporal fashion as TCCON before drawing any conclusions about comparative AGRs.

Response 3.42:

- We use the Global Carbon Budget and satellite data from Buchwitz et al., (2018) as the AGR references representative for global scales of CO_2 growth. These datasets are sampled by those researchers who generated the respective datasets and published these results. SAT and GCB are not necessarily sampled in the same way but as they represent the global growth, the agreement between them is high (r = 0.87). Therefore, we expect high agreement between the AGR references and (presumably) globally representative AGR_{TCCON}.
- AGR_{TCCON} estimates do exhibit high agreement with AGR references as shown *"AGR_{TCCON} strongly agrees with SAT (r = 0.83) and with GCB (r = 0.82) identically resembling SAT-to-GCB mutual agreement in global AGR reproduction (r = 0.83)"* except years 2008 and 2015.
- The issues from 2008 and 2015 are identified based on AGR-ENSO comparison shown on Figures 10 and 11. These are the years when probably ENSO-driven bias has influenced the accuracy of $TCCON_{AGR}$ estimates. As we stated, *"Despite it is challenging to outline the exact mechanism of this finding, if we assume 2008 and 2015 years were impacted by excessive or irregular sensitivity of some TCCON stations to short-term ENSO conditions."*
- As we are unable to explain the exact mechanism of the TCCON-to-reference bias exposure to ONI, we welcome any suggestions about the phenomenon of the AGR_{bias}-to-ONI agreement in MJJ (May-June-July) period (see Figures 10 and 11).

– line 416: From->"From a" This sentence is removed from the current version.

– line 420: this is not perfect correlation.

Response 3.43. Perfect correlation is everywhere reformulated to "excellent" if r = 0.98-0.99

– Figure 5: In the "Global" panel (top right side), there are TCCON data for 2006 and 2007 (gray bars), but there aren't any in the 3 other panels (neither in "NH", nor in "SH" or "Tropical"). Maybe emphasise that there must be a minimum 2(?) sites contributing within each region.

Response 3.44: The regional analysis is removed from the current revision due to lack of valuable information (for new research aim and objectives) for the current analysis and following several recommendations.

– line 434: variability->coverage or representation Reformulated

– line 436: dimensions? Do you mean on different scales? Yes. Reformulated

– line 438: errors-> differences OK

– line 439: it’s not clear how the time lag affects model AGR differences, unless there are differences in the models capturing this.

Response 3.45: True. We also did not notice this problem based on our data, so this assumption is not necessary. Sentence is reformulated, we mention only transport-model-driven errors.

– line 446: citations needed here

Response 3.46: To avoid misunderstanding, we reformulated the sentences by using word “reference data from satellite and global carbon budget” so a reader can understand we have in mind the reference data used in this article. The respective references are also added.

– line 454: again, not "perfect" agreement OK

– line 456: Oceans doesn’t need to be capitalized OK

– line 474-475: Not sure what this sentence means. Reformulated

– line 476: that->as OK

– line 477-481: lots of reporting of statistics without any need. This could be simplified.

Response 3.47: This part of paragraph is shortened.

– line 492: approves->supports OK

– line 493: not perfect OK

– line 498: "oppositely" - suggest replacing with an alternative word. OK

– line 501: at->in the OK

– line 504: "of by"-> "by" OK

– line 511: again, not perfect OK

– line 515/Table 2: Please add units for "Difference". OK, ‘ppm’ added

– Table 2 and lines 493-500: The median differences and correlation co-efficients are incosistent between the table and the text.

Response 3.48: Inconsistencies are corrected

– line 530: agreements->agreement OK

– line 538-540: revise as appropriate once earlier section is revisited Reformulated

– line 542: another misuse of "perfect", and in this case I wouldn't even call them near-perfect. Reformulated

– line 552: not perfect OK

– line 553: sentence needs revising ("on > 90%" doesn't make sense)

Response 3.49: In the new version, the formulations in these paragraphs are presented in different way and this reference is not required anymore.

– line 555: not perfect OK

– line 558: 0.02-0.03ppm->0.04 ppm? (c.f. line 494 and 495) OK

– line 566: "The results of this study have three vectors of implications." What does this mean? Why not just say there results have three implications? OK

– Conclusions: as noted earlier, "at second" and "at third" should be replaced by "secondly" and "thirdly" OK

– line 567: approves->confirms; old->existing OK

– supplement, line 8: Table S.2.2-> Table S.2.1 OK

– supplement, line 64, 85: subscript $_2$ OK

References

Lindenmaier, R., Dubey, M. K., Henderson, B. G., Butterfield, Z. T., Herman, J. R., Rahn, T., and Lee, S.-H.: Multiscale observations of CO2, 13CO2, and 265 pollutants at Four Corners for emission verification and attribution, Proceedings of the National Academy of Sciences,111, 8386–8391, https://doi.org/10.1073/pnas.1321883111, http://www.pnas.org/content/111/23/8386.abstract, 2014.

Lindqvist, H., O'Dell, C.W., Basu, S., Boesch, H., Chevallier, F., Deutscher, N., Feng, L., Fisher, B., Hase, F., Inoue, M., Kivi, R., Morino, I.,Palmer, P. I., Parker, R., Schneider, M., Sussmann, R., and Yoshida, Y.: Does GOSAT capture the true seasonal cycle of carbon dioxide?,Atmospheric Chemistry and Physics, 15, 13 023–13 040, https://doi.org/10.5194/acp-15-13023-2015, https://www.atmos-chem-phys.net/

270 15/13023/2015/, 2015.

Yuan, Y., Sussmann, R., Rettinger, M., Ries, L., Petermeier, H., and Menzel, A.: Comparison of Continuous In-Situ CO2 Measurements with Co-Located Column-Averaged XCO2 TCCON/Satellite Observations and CarbonTracker Model Over the Zugspitze Region, Remote Sensing, 11, 2981, https://doi.org/10.3390/rs11242981, http://dx.doi.org/10.3390/rs11242981, 2019.

Response 3.50: The suggested references were added to the manuscript.