**RESPONSE TO THE REVIEWER #2**

In their work "The consistency between observations (TCCON, surface measurements and satellites) and $CO_2$ models in reproducing global CO2 growth rate", submitted to ACP, the authors investigate the agreement of the atmospheric CO2 (annual) growth rates from several data sources: total column CO2 from the network of ground-based Fourier Transform Spectrometers (TCCON), inverse model estimates from Carbon-Tracker 2017 and CAMS, and reported growth rates from the Global Carbon Budget as well as total column $CO_2$ from satellites. While quantifying the atmospheric $CO_2$ growth rate using different data sources might be a topic suitable for ACP, the scientific questions that this paper attempts to address are not sufficiently clearly defined. Thus, the research presented in the paper seems to somewhat suffer from this throughout the manuscript. The concepts, ideas, tools and data in the paper are not novel, and the methods used for data analysis are not entirely valid or sufficiently described. I consider that the paper should not be accepted for publication in its current form. A thorough revision starting from re-formulating the research questions would be necessary. In what follows, I will point out more specifically my concerns and comments on the manuscript, focusing more on the early sections as they naturally affect the rest of the manuscript. Because the comments are quite extensive, I will exceptionally not list technical corrections nor specific comments related to the English language in my review.

Response 2.01: We thank the reviewer for the valuable suggestions based on which we have substantially improved the manuscript. Point-by-point responses are given below and marked by green color. Please note that the definition of GR, AGR and MGR terms are given in the article. We have followed the recommendation for thorough revision in the way proposed by the reviewer in the comment above. Foremost, we reformulated the research question and objectives to be more specific (also after many comments from TCCON community) as following.

- New research aim is *"Our study aims to assess the robustness of GR estimates from the observations of the Total Carbon Column Observing Network (TCCON) given the importance of the network to the global carbon monitoring and its expanding observational coverage."*
- 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". This objective is now supported by expanded evidences about daily, monthly, annual and seasonal stability of TCCON AGR estimates, objective*
- *(b) "to examine the* $AGR_{TCCON}$ *agreement with the existing* $CO_2$ *growth references and its sensitivity to external factors".*
- *(c) to assess the exposure of* $CO_2$ *growth estimates at each TCCON station to external factors*

Specific comments:

Lines 64-89: The main message of this paragraph is unclear to the reader: are we interested in the global or local CO2 growth rate in this paper? The authors start by highlighting the need for an accurate knowledge of the global $CO_2$ growth rate; however, their analysis focuses on both local and global analysis without a clear focus or guidance to what essentially is important in the results and why the analysis has been made.

Response 2.02: According to the reformulated research aim, we are interested in the global $CO_2$ growth reproduced by TCCON global observations. Please note that the introduction was restructured accordingly where the 3rd paragraph is now dedicated to the importance of TCCON-based $CO_2$ growth estimation.

On lines 97-104, the authors state the three objectives of their study. Regarding these objectives (a)-(c), (a) I don't think that aggregating all TCCON data would represent the global $CO_2$ growth rate because of several reasons: even though the TCCON is in principle a (near-)global network, the instruments are mainly located in North America and Europe. In addition, several of them are affected by local sources (e.g., Paris, Tsukuba, Pasadena) that make them interesting for the evaluation of satellite $CO_2$ retrievals in urban circumstances but maybe less representative of the global $CO_2$ background for the purposes of this study.

Response 2.03:

- We think that our study has provided several proofs about relative suitability of TCCON observations for calculating $CO_2$ global growth. Namely: satisfactory stability of $CO_2$ growth estimates despite high variability of data availability from station to station and from year to year, reasonable agreement with the global references, low sensitivity of TCCON-based $CO_2$ growth to the presence of urban sites in the aggregated signal.
- The latter finding about urban areas overlaps with the reviewer's concern about the influence of local sources (see the entire section 3.4.2 dedicated to this issue). The urban influence on interannual TCCON signal is expressed by the agreement rate of MGR estimates between TCCON (can be influenced by local source) and the models (are likely too coarse to capture local-source related $CO_2$ variability). 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 hypothesis about malignant role of these sites in global AGR signal is not supported 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."*
- Regarding the spatial limitations of the TCCON network, we have discovered the enhanced sensitivity of $AGR_{TCCON}$ estimates to ENSO late spring-early summer (MJJ) anomalies (Figures 10, 11). This may be an indication of irregular response of TCCON stations to ENSO anomalies where AGR of

some stations is more heavily influenced by these anomalies. All the above-mentioned speculations are added to the current version.

Regarding (b) and (c), the estimation of spatiotemporal inconsistencies between inverse models has already been carried out in a number of studies. It is not clear whether this study brings anything new to the discussion. In particular, I found the analysis regarding (c) a bit rushed and shallow, and lacking important references to earlier studies that consider either these particular models or regions of interest (e.g., Lindqvist et al., 2015; Palmer et al., 2019). The analysis focuses on the atmospheric CO2 growth rate and several times mentions a "permanent" growth of CO2, which may be misleading to readers, considering the seasonal cycle of CO2.

Response 2.04: We agree we did not manage to show that the estimation of spatiotemporal inconsistencies between the inverse models is a core of this study. As mentioned, we dedicated most efforts to the TCCON-related analysis in this version (objectives 'a' and 'b') while the model intercomparison now represents additional analysis (objective 'c').

- As the model analysis was shallow, we provided more details about spatio-temporal analysis between $CO_2$ models (agreement between the models and biomass burning, biosphere, fossil fuel and oceanic fluxes plus agreement between the models and the surface $CO_2$ data indicating how well the surface-governed growth is captured by models).
- The word "permanent" is avoided in this revision.
- Lindqvist et al., 2015 and Palmer et al., 2019 are added as references to this study.
- Moreover, Hannakaisa Lindqvist is included as one of the key coauthors in this revision.

Section 2.1.1 (TCCON): The TCCON data policy requires that the authors contact the TCCON PI's in the preparation phase of the paper in order to guarantee that the data are used and interpreted correctly and also to agree on a potential co-authorship in case the TCCON data have a central role in the manuscript. Since the TCCON PI's have already commented on this issue separately, I do not focus on this more. I do want to add, however, that several issues in the manuscript regarding the interpretation of the results at specific TCCON sites would have been clarified in the preparation process of the manuscript in case the TCCON PIs had been contacted for the work.

Response 2.05: We apologize for the oversight with the TCCON policy. We hope that the new version of the manuscript composed in closer coordination with the TCCON community with their comments applied (see the open access TCCON collective comment at the ACP page of our paper) is significantly improved.

Sections 2.1.2, 2.1.3, and 2.1.4: These sections lack plenty of relevant details, such as version numbers of several of the prior flux components and a more detailed description

of the satellite data, even though these data were cited ("CO2 observations from SCIAMACHY and GOSAT" is not sufficient).

Response 2.06: Sufficient details about all datasets have been added, please see the updated descriptions in the section 2.1 highlighted by yellow color.

Section 2.2.1: Description of the methodology is not sufficient for reproducing the results. For example, it is not clearly described how the gaps in the data are considered.

Response 2.07: Description of the methodology is complemented by additional information (Section 2.1.1) highlighted by yellow color. Since the data used for the AGR estimation as input considerably varies in temporal (daily, monthly, annual) and spatial (station-wise) scales, we devoted one objective of the study to investigating the role of input data characteristics in $AGR_{TCCON}$ estimation. Namely, "*to estimate the robustness of AGRTCCON due to the data sampling, measurement gaps or difference in time series across the sites."* The results from the data structure analysis are shown in sections 3.1 and 3.2.2. The most important expression of the data gaps in the final $AGR_{TCCON}$ estimates is described in the methodology now. They are calculated "*based on the station-wise variability using method similar to Buchwitz et al., (2018). Namely, average standard deviation across TCCON stations (AGR) multiplied on* $\checkmark N_{total}/N$ *factor where N – number of stations used,* $N_{total}$ *– total number of stations in TCCON analysis".*

Are there any criteria for including or excluding some of the TCCON sites?

Response 2.08:

- See section 3.1 where we tried to include as much stations as possible using the minimum daily threshold of '2' (please see the same section for explanation of this term).
- Note that to calculate $CO_2$ growth annual growth projected for monthly scales (MGR) for instance, of January 2015, one needs to have not only $XCO_2$ from January 2015, but also from January 2014. Since paired monthly estimates are required, we had to filter out many useful monthly estimates.
- As we have stated, *"there are not enough measurements to calculate a single MGR from MAN and IND stations regardless to the daily threshold (see Figure 2). At several other stations (AMY, BUR, JPL) there is insufficient number of MGRs (< 8) during the entire study period to calculate a single* $AGR_{TCCON}$*. Hence, 6 stations (FCO, MAN, IND, AMY, BUT and JPL) are not used in this study".* Here '8' is an arbitrary value that indirectly ensures that sites with too few MGRs to calculate even a single AGR are not used in the final analysis.
- Following the recommendation from one of the TCCON PIs, we have tested the role of AMY missing site in the global AGR (AMY has 3 MGRs at the softest daily threshold of '2'). The impact of these scarce MGR estimates on global AGR would be just within instrumental uncertainties (~0.1 ppm) for 2016 and 2017 years.

Exact methodologies should also be described for comparisons of model and TCCON data (e.g., spatiotemporal interpolation of the gridded model data, averaging kernel corrections etc.).

Response 2.09: We moved section the description about calculation of pressure-weighted $XCO_2$ using model simulations to the main body of the manuscript. See Equations 1, 2, please and the respective description. Simple description is also added about horizontal collocation of TCCON observation and grid cell of the model.

The results and discussion sections suffer from a very scattered analysis which is rich in details but not in content, and lacks focus. Correlation analysis is not sufficient in case of time series: for example, a phase difference in the time series would result in a relatively weak correlation but the reason for the weak correlation would not be identified. At least some representative cases of the $XCO_2$ time series should be presented.

Response 2.10: We tried to enrich the current revision results by

- Including few representative TCCON sites (Tsukuba, Park Falls, Garmisch) for $XCO_2$ vs MGR vs AGR detailed analysis (Section 3.2.1).
- Including additional analysis for Ascension measurement site (Figure S.2.6) where the modeled MGR values are seen as well.
- Regarding the AGR correlation analysis, when we use it for validating $AGR_{TCCON}$ against AGR references, we do not a phase difference between two different estimates of annual growth. We agree this can be the case for AGR-to-ENSO comparison, so the AGR-to-ENSO strength correlation analysis is provided with ± 1-year lag (Figure 9). Moreover, as the AGR-to-ENSO agreement can be driven by some seasonal components of ENSO (expressed as ONI index), we performed analysis of AGR versus every type of ONI (covering all possible 3-month periods during the year) as shown on Figure 10.

The discussions and conclusions drawn on the claimed "biomass burning regions" seem particularly rushed and would have been relatively straightforward to check by the authors because at least CarbonTracker 2017 provides the imposed fire fluxes as a separate data field.

Response 2.11: We incorporated the analysis of Carbon Tracker 2017 fire fluxes (as well as all other components). See the details at Figure 14 and the respective section please.