Review and feedback on Labzovskii et al from TCCON PIs

Nicholas M. Deutscher[1], Debra Wunch[2], Thorsten Warneke[3], Dietrich G. Feist[4,5], Coleen M. Roehl[6], Voltaire A. Velazco[1], Matthias Buschmann[3], Ralf Sussmann[7], Isamu Morino[8], Hirofumi Ohyama[8], Manvendra K. Dubey[9], Rigel Kivi[10], Frank Hase[11], Matthias Schneider[11], Yao Té[12], and David W. T. Griffith[1]

[1]Centre for Atmospheric Chemistry, School of Earth, Atmospheric and Life Sciences, Faculty of Science, Medicine and Health, University of Wollongong, Wollongong, NSW, 2522, Australia
[2]Department of Physics, University of Toronto, Toronto, Ontario, Canada
[3]Institute of Environmental Physics, University of Bremen, Bremen, Germany
[4]Lehrstuhl für Physik der Atmosphäre, Ludwig-Maximilians-Universität München, Munich, Germany
[5]Institut für Physik der Atmosphäre, Deutsches Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen, Germany
[6]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA
[7]IMK-IFU, Karlsruhe Institute of Technology (KIT), Garmisch-Partenkirchen, Germany
[8]Satellite Observation Center, National Institute for Environmental Studies, Tsukuba, Japan
[9]Earth and Environmental Sciences, Los Alamos National Laboratory, Los Alamos, NM, USA
[10]Space and Earth Observation Centre, Finnish Meteorological Institute, Sodankylä, Finland
[11]IMK-ASF, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[12]Laboratoire d'Etudes du Rayonnement et de la Matière en Astrophysique et Atmosphères (LERMA-IPSL), Sorbonne Université, CNRS, Observatoire de Paris, PSL Université, Paris, France

Correspondence: Nicholas M. Deutscher (ndeutsch@uow.edu.au)

The paper by Labzovskii et al examines the consistency of growth rates of CO_2 amongst different measurement datasets and models. There are an extensive range of measurements and models included in the analysis; however, we feel that a number of improvements should be made before the paper can be published in ACP. The following comments provide a number of suggestions about how that can be done, with an emphasis on the role and interpretation of TCCON data within the manuscript.

1 Major Comments

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

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

– 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 (AGR_{TCCON}) and its variability with models and satellite methods. One key finding is that the AGR_{TCCON} 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, recognising 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.

– 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.

– 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. Four Corners TCCON site was located near a power plant, which led to early morning plumes that increased xCO2 by up to 10ppm (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.
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.
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.
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 neighbouring 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.

- 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)CO_2
- As indicated in our other comment, please ensure the TCCON data DOIs are appropriately cited for the 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.
- 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.

2 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.

- line 9: remove "the" before "global warming"
- line10-11: "Despite atmospheric CO_2 growth rate had been considered as the well-known quantity"
  This wording doesn't work. Maybe "Despite the atmospheric CO_2 growth rate having been considered as a well-known quantity"
- line 13: Please correctly define TCCON as the "Total Carbon Column Observing Network"

- line 23-25: The structure of this sentence is confusing, and forms a double negative. It would perhaps be better to emphasise good agreement at 85% of stations.
- line 27: "perfect (r=0.99)" - not quite perfect
- line 29: in → a; also again, this is not "perfect".
- line 29: insert "a" before "spatial"
- line 33-36: sentence seems back to front. Missing 'a' before CO_2
- 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
- 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 CO_2 growth rate (GR)". There are many instances of this throughout the manuscript that we will not explicitly point out each time.
- 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").
- line 47: "uploaded to" - a better phrasing might be "reported in"
- line 48: at → "on a"
- line 50: than → that
- line 51: "constrains" can have a specific meaning in flux estimation, perhaps replace with "limits"
- line 61: insert "the" before terrestrial
- line 67: process → processes
- line 70: insert comma before which, remove "per se"
- line 71: ecosystem → ecosystems
- line 72: semiarid → semi-arid
- line 73: "Despite the regions" → "Despite the fact that the regions"
- line 74: evidences → evidence (and are → is)

- 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

- 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. 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.

- line 85: ratio → ratios

- line 88: remove semicolon

- line 96: add "the" before "Global Carbon Budget"

- line 100: "the TCCON" → "TCCON's"

- line 101: "approve" → "prove"

- line 102: "rest" → "remaining"

- line 105: "At second" → "Secondly" (similarly next line for "At third"

- line 108: "Sect.s" → "Sects"

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

- 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

- 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"

- line 120: O_2 (subscript)

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

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

- line 123: superscript $^{-1}$

- line 125: delete "XCO_2"; accuracy → uncertainty (2-sigma)
- 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
- line 127: "from World Meteorological Organization onboard" "traceable to the World Meteorological Organization scale"
- line 130: exist → existing
- line 131: "familiarize" → "familiarize themselves"
- line 133: "addresses to" → uses
- line 135-137: not clear what this sentence means - needs clarification
- 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.
- line 150: Insert "The" at the start of the sentence, and "is" before "being"
- line 151: "developed by the Institute of Pierre Simon Laplace (LSCE)"
  LSCE stands for "Laboratoire des Sciences du Climat et de l'Environnement"
- line 157: Sentence starting here needs clarification/rewording.
- line 160: insert "the" before biomass
- line 161: insert "the" before "fire module" and needs clarification on GFAS (GFAS emissions possibly, or "uses the GFAS")
- line 164-165: how did you judge that "this process did not cause serious error propagation"?
- line 175: cover → coverage
- line 177: "CTA"
  What does this stand for? Replace with "CT"?
- 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.
- line 178: "GFED" is already defined in line 161
- line 188, 221, 224: subscript $_2$
- line 191: "being updated at" → "is updated on"

- line 192: why "Obviously"? Suggest removing this word.
- line 195: on → at; insert "The" before GCB
- line 196: The correct journal name is Earth System Science Data
- 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)
- line 201: Should be "Greenhouse gases Observing SATellite"
- line 210: is there also a neutral classification (0)?
- line 226, line 577: Piao et al., 2019 –> Piao et al., 2020
- 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
- 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.
- 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.
- 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 ?)"
- 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.
- line 273 "a study period" → "the study period"
- line 275: r = 0.61 for TCCON-SAT, r = 0.49 for TCCON-GCB In the Abstract and Discussion sections, the correlation co-efficients are described as being 0.75 for TCCON-SAT and 0.68 for TCCON-GCB. Which is correct?
- line 278: "right below" → "directly below the"
- 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.
- line 283: approve? Not the correct word here
- line 287: "(right axis of Fig. 1, panel b)" → "(right axis of Fig. 2, panel b)"
- line 291: "to data" → "with data"

- line 292: remove comma
- line 295: "Besides 15-year pattern" - not sure what you mean?
- line 298: "yield to" → "yield a"
- 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.
- 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.
- 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!
- line 319: "be risen" → "arise"
- line 320 "(based on one-time TCCON observations)"
  Is it one measurement of TCCON or is it a daily mean measurement of TCCON?
- line 321-322: Not sure what this sentence means - clarify.
- line 329: "different thresholds we mentioned above" → "the different thresholds mentioned above"
- 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.
- line 347: such as 30 points
- line 349: approves → proves (?)
- line 363: versus → with
- line 367: add "the" before "correlation"
- line 378: plain → simplistic (assuming this is what you mean)

- 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 hypothesise, surely the neighbouring sites (Rikubetsu, Saga) should exhibit similar behaviour.

- 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.

- line 397: add "themself" after familiarize

- 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 co-efficients on the plot.

- line 415: on → by?
  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.

- line 416: From → "From a"

- line 420: this is not perfect correlation.

- 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.

- line 434: variability → coverage or representation

- line 436: dimensions? Do you mean on different scales?

- line 438: errors → differences

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

- line 446: citations needed here

- line 454: again, not "perfect" agreement

- line 456: Oceans doesn't need to be capitalised

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

- line 476: that → as
- line 477-481: lots of reporting of statistics without any need. This could be simplified.
- line 492: approves → supports
- line 493: not perfect
- line 498: "oppositely" - suggest replacing with an alternative word.
- line 501: at → in the
- line 504: "of by" → "by"
- line 511: again, not perfect
- line 515/Table 2: Please add units for "Difference".
- Table 2 and lines 493-500: The median differences and correlation co-efficients are incosistent between the table and the text.
- line 530: agreements → agreement
- line 538-540: revise as appropriate once earlier section is revisited
- line 542: another misuse of "perfect", and in this case I wouldn't even call them near-perfect.
- line 552: not perfect
- line 553: sentence needs revising ("on > 90%" doesn't make sense)
- line 555: not perfect
- line 558: 0.02-0.03ppm → 0.04 ppm? (c.f. line 494 and 495)
- 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?
- Conclusions: as noted earlier, "at second" and "at third" should be replaced by "secondly" and "thirdly"
- line 567: approves → confirms; old → existing
- supplement, line 8: Table S.2.2 → Table S.2.1
- supplement, line 64, 85: subscript $_2$

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 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/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.