Review of Metz et al.: Seasonal and inter-annual variability of carbon fluxes in southern Africa seen by GOSAT, submitted to Biogeosciences

14 Oct 2024

Review Summary

The study by Metz et al. starts from satellite-retrieved XCO2 data analysis over a large southern African region dominated by semi-arid savannas and grasslands. They find a larger seasonal variability of XCO2 from GOSAT and OCO-2 compared to an ensemble of in-situ-optimized inverse models resampled at GOSAT soundings. They further point out that the discrepancy originates from the biospheric prior fluxes that control the optimized biospheric fluxes due to the lack of observational constraints. An understanding of the processes that affect the CO2 fluxes and their variability is gained through a comparison of the TRENDY v9 ensemble of Dynamic Global Vegetation Models (DGVMs).

The paper is well in the scope of Biogeosciences. The paper points out large discrepancies in DGVMs over semi-arid regions in southern Africa, and presents how satellite data can be used to evaluate these models. While the results of the paper do not include completely novel findings (like the authors already point out with appropriate references in the paper), the results are in agreement with previous studies carried out using other approaches, and should therefore be interesting to the readers of the journal. The paper is particularly well written, and the methods and results are clearly presented. I also like how the paper advances from one result to the next one in a way that is pleasant to follow. I recommend publishing the paper in Biogeosciences after the authors have considered the constructive comments detailed below. I have also gone through the comments from Reviewer 1 and try not to repeat their message, although some of my comments might be of similar nature.

Major comments

1) I don't think the title of the paper is fully descriptive of its contents. Could the title be for instance along these lines: "The potential of CO2 satellite observations to inform and evaluate modelled carbon dioxide fluxes in Southern Africa" or something similar. At least, I would propose to change "carbon" to "carbon dioxide" or "CO2" because methane and CO are not discussed in the paper.

2) The study region is very large and the CO2 fluxes can be quite heterogenous within the study region (for example, on NOAA's website one can see that CarbonTracker concludes part of the study region to be a net source and part a net sink of CO2; https://gml.noaa.gov/ccgg/carbontracker/fluxmaps.php?region=afr&average=annual#imagetable and even though CT is not the best-performing model in the

region as the authors point out, this underlines the heterogeneity of the region and makes one wonder how much you can average over without losing important information). I think that the averaged XCO2 time series (Fig. 2) is to some extent dependent on the sampling. The number of GOSAT observations likely varies from only few tens to some 1000-2000 per month over this region, depending on the season. It would be good to show the time series of the number of GOSAT observations (both products) for example added in Fig. 2 and also a map of the spatial sampling density for example around the XCO2 minimum and XCO2 maximum (this could be in the Appendix), and discuss how the sampling affects the results of the concentration-based analysis and the inverse modelling results.

Minor comments

Abstract: I suggest highlighting the discovered role of grassland carbon uptake for IAV which I consider as a key result of the work. It would also be helpful to give key numbers of TgC so that the reader can contrast those more easily.
Line 46: Spell Orbiting Carbon Observatory -2 with a hyphen
Lines 56-57: CO2 exchange fluxes → CO2 exchange
Line 57: to flux estimates → to those (to avoid repetition)
Line 69: Valentinti → Valentini (check!)
Line 69: grass land → grassland
Figure 1: This is a massive study region. To make this clearer to the reader, I suggest adding a scale bar representing for example 200 or 500 km or similar. In addition, please add the locations of the Gobabeb COCCON and the Kruger National Park measurement with pins or equivalent symbols.
Line 76: I believe that the description of the timing of the fire season applies to the study region in question and not the whole Africa. Please specify in the text.
Section 2.2: Did you also include satellite data over the ocean in your analysis?
Line 82: column-average → column-averaged
Line 85: How different is RemoTeC v2.4.0 compared to RemoTeC v2.3.8? Please summarise the key developments after v2.3.8 in the text.
Lines 87-90: I understand that both products have been used as two separate products (e.g., apply their own quality filtering and potential bias corrections). Did you compare ACOS and RemoTeC by considering only the exact same GOSAT soundings? If so, what did you learn about the product XCO2 differences? While this question may not be in the core of the paper, I think it is of high interest to learn as much as possible about the satellite observations' differences in particular over these regions that are poorly sampled by any other measurements. In addition, the differences in the satellite products can result in flux differences of several tens TgC per month for this region (Fig. 3), which is significant and interesting. Depending on the authors' view, this might even merit a sentence in the Conclusions to address the need of validation measurements in this region to be able to reconcile residual differences in the satellite products.
Lines 87-90: In Fig. 2, we are seeing differences in the sub-ppm scale. It would be helpful to give guidance (with references) to the (global) accuracy and precision of

these GOSAT products in this part of the text to help the reader understand to what extent we can trust the satellite observations.
Line 91: I think it would be more descriptive to say "evaluation" instead of "validation".
Line 94: The COCCON data were not co-located with the satellite observations, right? Did you do any filtering to the data, for example in the early or late hours with large solar zenith angles? Please specify in the text.
Sect. 2.3.1: The three inversion models also use different data assimilation schemes, right (4DVar vs. enKf)? This could be specified in the text.
Sect. 2.3.1.: Did you assimilate also GOSAT observations over oceans? Please specify.
Sect. 2.3.1.: Question, out of curiosity: did you also consider a GOSAT-only inversion with no in-situ data?
Line 113 (and elsewhere): please use \citep[TM5-4DVar,][]{Basu et al reference} LaTeX command to get the rid of the extra parentheses within parentheses (if working with LaTeX).
Lines 141-144: Did you apply the averaging-kernel correction when comparing the cosampled model concentrations to satellite data? The effect is not likely to be large but it might be non-negligible since the differences in question are mostly in the sub-ppm scale.
Line 156: Does the FLUXCOM have a product version that could be referenced here? We have recently worked with a FLUXCOM product that had a strong positive bias in CO_2 fluxes over both Northern and Southern Africa. This seems to not be the case with the product in the paper. In fact, if I interpret Fig. 3 correctly, the FLUXCOM NEE would give out either close to zero or negative values without GFED fire emissions, correct?
Line 173: Methods → Data and Methods
Lines 176-178: This reminds me of similar GOSAT-to-model discrepancies that were pointed out by Lindqvist et al. (ACP, 2015; https://acp.copernicus.org/articles/15/13023/2015/) and their reasoning, albeit with much earlier versions of the data and models.
Figure 2: Please consider adding the monthly data density or, if the result is too crowded, presenting the data density time series using other means.
Figure 2: What does the CS in the subscript of the "Inverse models" stand for?
Line 190: e. g. → e.g.,
Sect. 3.2: I think it is interesting and worth pointing out how you can already "see" the impact of the fluxes in the detrended concentrations (Fig. 2). The well-aligned timings also suggest that the concentration maxima and minima are driven by the fluxes from this region as opposed to transport from elsewhere. This suggests that satellite data can already in its concentration form have potential and be useful for an evaluation of models in a region (even without inverse modelling!).
Figure 3 (a): the legend is partly on top of the plotted data. Please revise such that the data are fully visible. This comment applies to some subsequent figures as well.
Line 204: remove "methods"
Line 244: a → an
Line 253: processed based → process-based
Sect. 3.3.: I understand that if the focus of the paper is to learn about the processes affecting the large-scale CO_2 exchange in this region, it is meaningful to focus on the DGVMs that agree with the GOSAT inversions. However, as also Reviewer 1 points out, it seems like a missed opportunity to not discuss the models that agree the worst. I also

wonder how generalizable these findings might be for other semi-arid regions in the world; one good example already published by the authors about the Australian fluxes (Metz et al., 2023). I get the feeling that the authors have already learned much more from individual model performance than what they are writing in the paper. I know it is not a preferred task for anyone to point out flaws in someone else's product but this is sometimes necessary to enhance openness that might further advance science.

Figure 6: The shading mentioned in the caption is not visible (at least in my printed version). The legend covers part of the data, please revise the figure. Why is GOME-2 SIF used instead of GOSAT SIF or OCO-2 SIF? I wonder how different these would be.
Lines 278-279: CAMS cannot be identified from the spread of models in Figs. 2-3.
Lines 283-284: You could add a very recent reference in your discussion of fire emissions in the regions. This one also seems to suggest underestimated fire emissions. Van der Velde et al. (2024):
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2023GL106122
Lines 300 and 304: I think the correct spelling is La Niña and El Niño
Line 302: "enhanced surface near soil moisture" this sentence might be missing a word?
Sect. 3.4 title: I'm not sure what is meant by gross fluxes. Could you just say fluxes?
Line 321: remove "fluxes"
Line 344: Suggest to rewrite the sentence starting "A by 1-2 months…" as it is not easy to read.
Lines 359-362: To me, carbon uptake by grasslands dominating the IAV of the entire southern African fluxes sounds like an important finding that should be brought to the Abstract and perhaps even contrasted in magnitude against fire emissions to the reader so that the magnitude is easier to grasp. At least to me this finding highlights the importance of an improved understanding (and modelling) of the sensitivity of grasslands.
Lines 359-362: This is an example where the sampling of the satellite data for GOSAT inversions might end up affecting the resulting reasoning (see my Major comment #2).
Line 374: rain induces → should this read rain-induced?
Figure A1 caption: und → and
Figure A1 legend: what does "CS" stand for? Also, the legend is covering part of the data, please revise.
Figure A3: it seems that months 10 and 11 are particularly challenging. Could you please add a bit of discussion related to this in the text?
Figure A8: Just a comment: this amount of precipitation is likely to affect the number of satellite observations of the region. Please see my Major comment #2.
Figure A9: Are these measurements assimilated in any of the in situ optimized inversion models discussed in the paper? The time span would match part of the time period discussed in the paper.