

1 General remarks

The manuscript by Jacobs et al. describes updates to the OCO-2 ACOS v11 product with a focus on the improvements gained by using an updated digital elevation model. They compare various DEMs with a focus on the high latitudes and investigate the improvements gained by comparison to TCCON and the impact on flux estimates. The manuscript is overall well written, however some areas could benefit from additional information.

Data description of OCO-2 retrievals

In the manuscript measurements taken in target, land nadir and land glint mode are mentioned and used. Please provide some explanation on the properties and differences between measurements taken in these modes. This would be especially important in context of the results presented in Table 6. Why was the initial bias in v10 lower in LNLG measurements compared to target mode measurements?

Inversion results

In section 4.6 the impact of the DEM on flux estimates is discussed by comparing both the NASADEM+ and Copernicus DEM fluxes to the OCODEM fluxes. To me it seems that no significant difference in zonal mean fluxes are visible when using the Copernicus DEM. When using NASADEM+ the difference in zonal mean flux is however clearly visible. I think this is not sufficiently discussed in the paper and caused some confusion for me.

1. In the abstract and conclusion the stronger differences when using NASADEM+ are highlighted. In v11.1 the NASADEM+ is however replaced by the Copernicus DEM which shows no such differences. This seems confusing to me and this highlighting needs more justification (or it should be removed). At first glance this can be interpreted as an exciting result gained from using a new DEM (compared to v10). But I suppose this is meant to discourage the usage of v11 in comparison to v11.1?
2. In Figure 13 panels (e) and (h) look very similar, while (f) and (i) (or (d) and (g)) do not. Why is this the case?
3. In Figure 13 the maps are shown with a $4^\circ \times 5^\circ$ resolution, which is coarser than the resolution used when comparing DEMs. Why is this the case? Is this the resolution of the flux inversion? This should be clarified and also mentioned in section 3.4.

2 Specific comments

- Page 1 Line 10: Why highlight the flux differences when using NASA-DEM+, when you decide on using the Copernicus DEM in the updated version (which does not show significant flux differences)?
- Page 2 Line 24: Could you shortly specify what kind of flaws?
- Page 2 Line 26: Please specify what kind of effects the fixed surface pressure would have. Was the retrieval with a fixed surface pressure informed by an updated DEM tested?
- Page 3 Line 1: It would be helpful to provide a brief summary of the results from Hachmeister et al. (2022).
- Page 3 Line 22: Is this relevant for you comparisons between different DEMs for the XCO₂ retrieval? I don't think you follow up on this later in the manuscript.
- Page 4 Line 31: Why these two gases in particular? Would other trace gases be similarly affected?
- Page 5 Table 1: I am not sure whether this table is necessary since it contains no additional information.
- Page 6 Line 24 – 25: To me it is not clear what motivates these formulas. Where does the 0.016 come from and what is meant by: “co₂_ratio_bias and h₂o_ratio_bias represent the piecewise linear fits”?
- Page 8 Line 1-2: How was the averaging performed? What is the order of magnitude of these artifacts?
- Page 9 Line 1: Which version of the Copernicus DEM is used here? It would be more precise to use the abbreviation GLO-30 or GLO-90 for the 30m or 90m version respectively instead of “Copernicus DEM”.
- Page 9 Line 12: While Hachmeister et al. (2022) mention that the Copernicus DEM is used in the updated version of the XCH₄ data product, their analysis is based on comparisons between GMTED2010 and ICESat-2 data. Schneising et al. (2023) describe the updated XCH₄ data. This should be clarified.
- Page 9 Line 18: I do not understand what is meant by the word “fidelity” in this context.
- Page 9 Line 23: Please specify “most regions”.
- Page 9 Line 27: Please explain the abbreviation “TIN”.
- Page 10 Line 7: What resolution is used your work?

- Page 11 Line 25: What is the motivation for this collocation criterion?
- Page 12 Table 3 & 4: It might be a good idea to combine these tables since they have the same column names.
- Page 13 Table 5: What is the reason for this selection of TCCON sites? It seems that only a small number of sites are left out from the analysis. It seems to me that especially Ny-Ålesund should be included in the analysis, since (a) the largest differences between DEMs lie in the high latitudes and (b) the topography around the Ny-Ålesund site is mountainous.
- Page 34 Figure 4: Why is there a gap in the middle of subfigure (b)? It should be explained where this is coming from.