Interactive discussion on "Four years of global carbon cycle observed from OCO-2 version 9 and in situ data, and comparison to OCO-2 v7" by Peiro et al., (2021).

The manuscript written by Peiro et al., (2021) is an update of the study made by Crowell et al., (2019), et al., (2019), who assessed the annual and monthly ensemble mean flux derived by nine different global inversions using the assimilation of OCO-2 satellite data (version 7) and in situ data for the period 2015-2016. Helene et al., 2019 used the findings found in Crowell et al. (2019) and compared them with the ensemble median flux of 10 different global inversions based on the assimilation of OCO-2 (version 9) land nadir and land glint (LNLG) data. Further, the authors also analysed fluxes for a more extended period 2015-2018, using both in situ and OCO-2 satellite data separately. In general, the manuscript is well-written, and the discussion and conclusions are interesting. However, some parts in the methodology and results sections need revision and clarifications. I would recommend this paper for publication once the authors have addressed all the questions described below.

General comments:

Line 15. "However, the lack of data in the tropics limits our conclusions and the estimation of carbon emissions over tropical Africa require further analysis". It would be good If the author can provide spatial maps of LNLG OCO-2 across the world for the studying years to see how bad the coverage is in Tropics. (Maybe show these maps in the appendix of the manuscript).

Line 137. "This bias correction in v9 allows a more uniformly distribution of XCO2 over regions of interest, decreasing the standard deviation to 0.74 ppm compared to v8, which was of 1.35 ppm." Please indicate that this reduction was found in TCCON Lauder, New Zealand (Fig.13).

Line 147-151 Interestingly, the authors accounted for correlations between the model-data mismatch error for the 10 second average of OCO2 soundings. However, the description of how it was calculated is not clear. It would be great if the author could provide more details in this section of the manuscript.

Line 159. "Both resulting uncertainty is then passed through the formula to account for error correlation." Which formula? It is hard to understand what the author did to calculate the observation uncertainties. Please provide more details in this section.

Line 160. "An additional term is added in quadrature to account for transport model errors. This model error term is computed from the difference between the CO2 concentrations computed by the TM5 and GEOS-Chem models when both are driven by the same realistic surface CO2 fluxes, after the annual mean difference filed is subtracted off".

The author referenced Basu et al., (2018) when they described the transport model error. In Basu et al., (2018), the model part error was calculated by considering the difference between a suite of inverse models optimised with in situ data and OCO-2 retrievals? Did the author do something similar here? Please clarify.

Line 240. For LNLGv9 inversions, the available 10s OCO-2 retrievals were averaged and compared to TCCON observations". I think the author means prior and posterior XCO2 simulated retrievals were averaged and compared to TCCON?

Line 257-258. Please indicates that Figure 5 represent only the median annual flux and the monthly median flux. In Cronwell et al., 2019, the authors discussed the annual and monthly mean flux, not the median. Could you indicate why you decided to compare the median and not the mean of the flux?

Line 261-262: “However, the peak sinks during the Northern Hemisphere growing season (from May through September) are slightly larger with OCO-2 v7 than with OCO-2 v9”. Something is missing here. In this section, you discuss global flux estimates, and at the same time, the author discusses fluxes in the Northern hemisphere. I cannot see the Northern Hemisphere Figure for this section.

The author includes Lofi inversion in the paper. However, there are no comments about their results. For example, Lofi seems to agree with the prior median estimate over the Northern Extra Tropics but not with IS and LNLGv9 assimilated flux.

Line 310- 313. “In the Southern Extra-Tropics, the authors mention that the sink flux estimate with OCO-2v9 is stronger than estimates made with IS, and v7, mainly because the ensemble spread with v9 is larger than with v7.” I am a bit surprised that fluxes assimilated with v9 have a larger spread than v7. OCO-2 v9 data supposed to have lower biases than v7. It seems that the error in the transport model might be the consequence of the large spread in the inversion with v9.

Line 314-318. Over the Tropics, the authors mention that posterior flux estimate (LNLG) is quite different to the prior estimate (Fig 6.c), but not comments are provided. How reliable can be the posterior estimate over the Tropics knowing that OCO-2 retrievals can be biased due to cloud coverage during the wet season and aerosol from biomass burning during the dry season, as the author mentioned?

Could you explain a bit more about the dipole in northern Africa? You mentioned several studies, but a better explanation is needed.

Fig 8 (NH Tropics) and (SH Tropics) LoFi monthly seasonally seem to be offset? Could the author explain why LoFi might not be capturing the seasonally over these latitudes bands?

Line 412. Alternatively, the OCO-2 inversions could be dominated by savanna seasonality (Baker et al., 2021a in prep). I cannot find (Baker et al., 2021a) in the reference list, and the reference cited (Baker et al., 2021) does not mention anything about the savanna seasonality but how to calculate the error correlations in OCO-2 data. Please provide the correct reference.

Line 426-. "Over the southern hemisphere, a large underestimation of the ensemble mean of LNLGv9 appears compared to the observations". I am surprised with these findings; I would have thought that the spatial coverage that OCO-2 data in the southern hemisphere would improve the results compared to IS. I am also surprised by the significant difference between LNLGv9 and IS biases in SH. Do you know how large are LNLGv9 OCO-2 biases in SH compared to v7? You mentioned in the introduction that biases in LNLG v9 were reduced considerably compared to v8 and v7. Is it possible to provide a Table in the appendix with no normalized bias?

Line 466. It is possible then that this excess of concentration, in both experiments, reflects the initial conditions of the inversion. Is there any way to test this?

Fig 13 and Fig 14. I don't understand why the author normalized the biases. It is clear that RMSE is large at some latitudinal bands and TCCON sites, suggesting that raw biases might also be high. I would also consider adding this information to the manuscript.

Line 471-474. "Compared to the evaluation of ISv7 in the study of Crowell et al. (2019), ISv9 and LNLGv9 biases are closer to each other (in the v7 MIP, the LNv7 were biased high compared to ISv7). Additionally, the OCO-2 biases have decreased (to values between -1.0 and 1.0 ppm) with v9 compared to v7, where biases ranged between -1.5 and 1.5 ppm (Crowell et al., 2019)". I don't understand why the authors say that OCO-2 biases have decreased compared to Crowell et al., 2019. Did Crowell et al., 2019 normalize the bias? If not, I don't quite understand the comparison.

Line 474-476. "…to the accuracy of TCCON retrievals over these regions (Crowell et al., 2019). I think that biases might likely be associated with biases to satellite observations than TCCON bias. Besides, I dont think that Cronwell et al., 2019 is not a good reference for talking about the accuracy of TCCON. XCO2 TCCON retrievals can contain airmass-dependent biases, which are corrected using the method described by Wunch et at.. Just wondering how wrong could be this correction to cause the posterior concentration biases seen here.

Line 476. Could the authors provide the location of the TCCON sites in a map instead of Table 3? It would be better for the reader to see where Caltech, Saga and Tenerife TCCON sites are located. I had to look at their locations from other manuscript.

The Caltech, Saga, and Tenerife sites show large underestimations in the IS and v9 results across all models. Isn't it bad to find a large underestimation at Caltech for the reliability of the fluxes estimated in North America? Is the location of Caltech a coastal site in the model? If so, it might strongly affect by ocean fluxes where not OCO-2 observations were assimilated.

Line 480. Another possible explanation of the underestimation observed over Saga and Izaga is that these small islands are strongly influenced by ocean fluxes, where the assumed uncertainties are small compared to land.

The biases in v9 have decreased for Ascension Island compared to v7, where the biases were around 1.0 ppm for LN and LG. Are these results compared to TCCON biases presented in Cronwell 2019? If so, I don't understand why the authors compare standardized bias against not standardized biases?

Line 495. "Transport model uncertainty is not expected to have changed dramatically since v7. This suggests that the reduction in the ensemble spread is likely related to a decrease in OCO-2 retrievals errors in v9 compared to v7". Are the modellers that participate in OCO-2 MIP have trying to consider improvement in the transport modelling (at the surface)?

Line 525-526. "When they compared..." who is they? I think the reference is missed here.

In the discussion section, I think it is also important that the author indicates why the assimilating OCO2 LNLGv9 data over the Tropics shows a stronger seasonality than prior fluxes. For the reader, it would be good to know why prior fluxes derived from ORCHIDDE or CARDAMON are not capturing well the seasonality. What could be wrong with these models?

Line 552 -554 "…with a slight negative bias in the v9 OCO-2 data for almost all latitudes, particularly

in the Southern Hemisphere and the tropics, where few evaluation data are available". Here the author mentions "slight negative bias in the v9 OCO-2 data", however in lines 426-427, they write, "Over the southern hemisphere, a large underestimation of the ensemble mean of LNLGv9 appears compared to the observations". This seems a bit contradictory.

Editorial and minor comments:

Title: Four years of global carbon cycle observed from OCO-2 version 9 and in situ data, and comparison to OCO-2 v7. As a personal opinion, I would write version 7 instead of v7 to be consistent with "version 9".

Line 121. "Inversions assimilating OCO-2 ocean retrievals produced unrealistic results with annual global ocean sinks higher of 2.6 ± 0.5 GtC.yr−1 compared to the state-of-the-art estimated in Le Quéré et al. (2018)." Could you quote the ocean sink estimated in Le Quéré et al. (2018).

Line 136. (Kiel et al. (2019)). Remove parenthesis to 2019.

Line 151. "Details of the form and derivation of these average uncertainties may be found in the 'constant correlation' section of Baker et al. (2021).", Please be consistent with the quotation mark "constant correlation".

Line 137. "This bias correction in v9 allows a more uniformly. Replace the word uniformly by uniform.

Line 184. There is an extra point. Please remove.

Line 241-242 Figure3. Please add the y-axis description (right legend).

Line 311. "However, in contrast to NHExt". In this line, you need to define define NHEx.

Line 305: The Southern extra-tropics. Capital letter?.

Line 306. Correct the word "signifcaintly" to significantly.

Line 398. For the monthly emissions of the Southern Tropical regions (Fig. 10), we can see the strong impact of El Niño in fall 2015 over Southern Tropical Asia in the larger emissions (with a maximum of 0.35 ±0.01 PgC/yr) given by all the inversions compared to the rest of the period. Please re-write, it is hard to understand.

Line 400. This large fall 2015 mainly come from Indonesian fires. I would write: The large emissions from Southern Tropical Asia (Fig.10.f) primarily come from Indonesia fires.

I would also remove the cyan and green lines from all plots (Fig.5 to Fig.10) that show monthly median fluxes for 2018-2019. If there is no data there, I don't think it is needed to show that the median is zero.

Line 605. … is about 0.5 GtC.yr−1.Land (need space).