Reply to Reviewer #1

The authors designed a two-step inversion approach to study the impact of fires on the inversion estimation of net carbon sources and sinks. In the first step, CO fire emissions are constrained with MOPITT CO data, and these optimized emissions are translated into CO2 emissions using biome specific emission factors. Subsequently, these optimized CO2 fire emissions are used as input in a second inversion step - alongside a rebalanced prior NEE that fits the global atmospheric growth rate - to constrain regional NEE with OCO-2 and in situ CO2 data. Despite much information is provided and carefully investigated, I do have a number of concerns about the results presented. The main text also requires additional editing and a thorough proofread to improve the general readability. In particular, I would like to see a shortened results section with a focus on the main findings and a smaller number of figures. The lengthy descriptions in each section distracts from the main points of this original work. Perhaps some of detailed descriptions, figures and comparisons can be moved to a supplementary document. The points below must be addressed before the paper can be accepted for publication in ACP.

We thanks Reviewer #1 for taking the time to review our manuscript and for the comments. We answered below the comments with information on page and line numbers that have been changed in the manuscript when necessary. Red is for suppression or modification, while blue is for sentences added in the manuscript.

Main concerns:

1. Reported optimized fire CO emissions:

1. After reading the first part of the paper, a number of things struck me regarding the CO emissions. Figure 5 provides an overview of the prior (GFED4s) and posterior CO emissions. I find it curious why the Northern Tropical and Southern Tropical African fires are scaled down so much. This contradicts other recent studies of African fires. Zheng et al. (2018), which you have included in the reference list, did a similar kind of inversion for the African continent with MOPITT CO data. For the period 2005-2016, they found more or less similar emissions for Northern Africa compared to GFED4s. But for southern Africa, they found that GFED4s underestimated the emissions by about 62%. Your results show the opposite; GFED4s overestimate emissions for Africa. Of course, there are differences in the time period studied between the two papers, but nevertheless it looks surprising. I think that this difference in outcome should be noted in the paper and, if possible, also explained.

In our manuscript, we indeed found that MOPITT CO data assimilated in TM5 has lower annual emissions for tropical Africa than observed with the fire prior. But posterior emissions are superior to the priors for the temperate regions of Africa. We only assimilated MOPITT CO v8 TIR-NIR data in our CO inversion using the model TM5. Our period of study is from 2015 through 2018, corresponding to 4 years of analysis. In Zheng et al., (2018) paper, they indeed looked at a different period: 2005 through 2016 corresponding to 11 years of analysis. Additionally, they did not assimilate the same data. They constrained their inversions with MOPITT CO v7 data as well as OMI CH2O (Ozone Monitoring Instrument, formaldehyde) and in-situ measurements from WDCGG (World Data Center for Greenhouse Gases) of CH4 and MCF (methyl chloroform) data. By adding these in-situ measurements in the inversions, they particularly help constrain OH, but also the photochemical reaction (source) of CO. Their chemistry is then different to the one used in TM5. They consequently have a multi-species atmospheric Bayesian inversion approach to get a well constrained global CO budget. This CO estimates has been shown to be higher than other inversions. In Zheng et al., 2019, they compared their CO biomass burning emissions with their multi-species estimation to nine previous studies using different inversion systems (Fig.10 of Zheng et al., (2019)). The use of multi-species inversion increases the annual average emissions of CO biomass burning at global scale by 20-37% compared to other inversions. Even if we used an updated version of MOPITT data compared to this study, we did not assimilate in-situ data and our chemistry is not as well performed as they used. As a reminder, the chemistry used in TM5 is based on climatology emissions (CO production based on a full TM4 chemistry of 2006). The model, transport, as well as assimilation algorithm used in Zheng et al., (2018, 2019) are also different to those used in TM5. It is then difficult to evaluate our results to their results.

Additionally, the figure below gives confidence that our posterior CO emissions fit the MOPITT data over Africa.

Finally, it is important to mention that Zheng et al., (2018, 2019), do not provide any evaluation against independent data. Therefore, there is no result proving that their results are more accurate or more trustworthy than our results.

2. Is this African signal driven purely by MOPITT CO? If so, can you demonstrate this by showing world maps with annual mean concentrations of prior CO (from GFED4s), posterior CO (from optimized fires) and MOPITT CO?

We thank the reviewer for this comment. We evaluated below the inversion's framework to the MOPITT measurement. We found, indeed, that the lower signal in our posterior is driven by low MOPITT concentrations over Africa. The figure below has been added to the supplementary information consequently and reference to this figure has been added in the section 3.1.

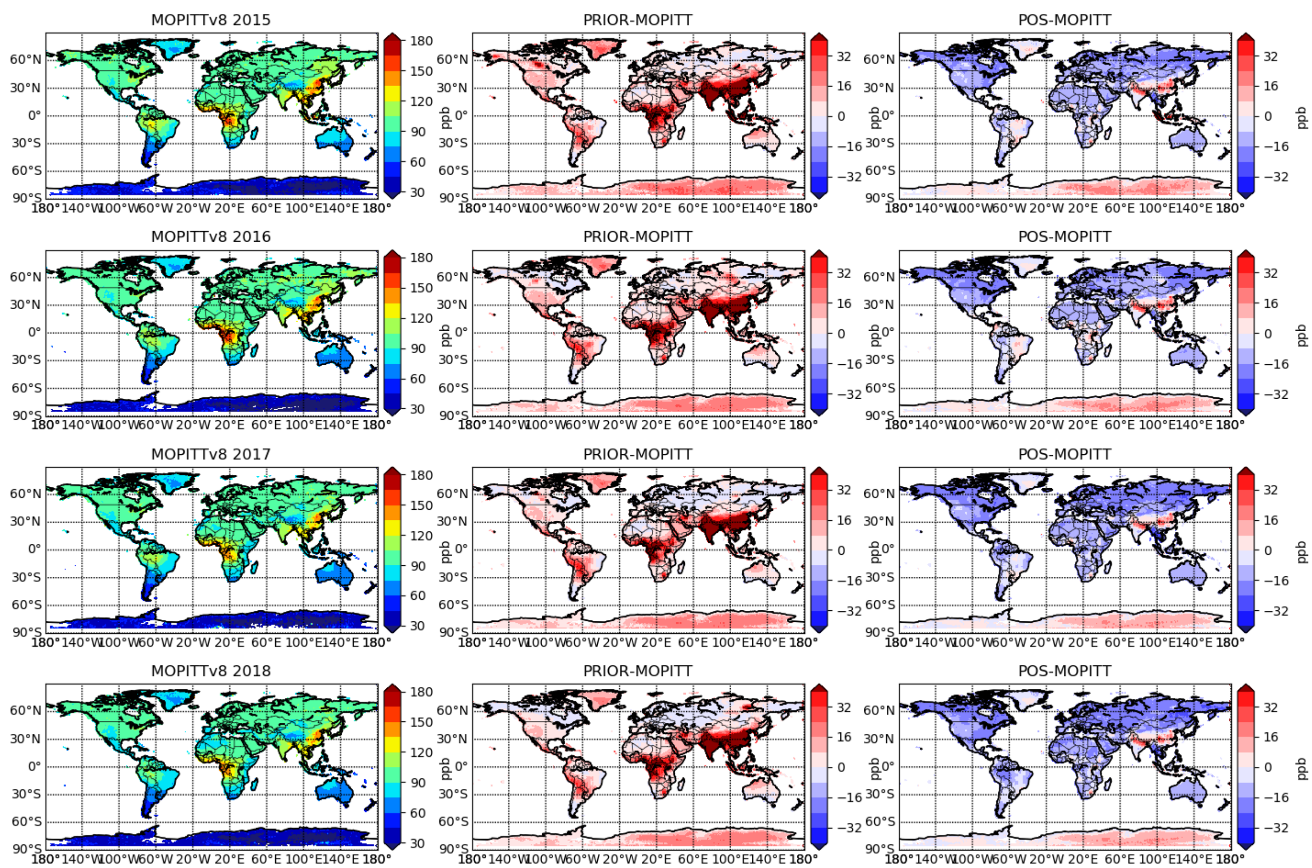

Figure S1b. Spatial distributions of the CO total column (XCO). Left column : distribution of annual mean XCO of MOPITTv8 retrieval. Center column : Mean annual difference between the prior simulation and MOPITT. Right column : Mean annual difference between the posterior simulation and MOPITT. From top to bottom are the annual mean from 2015 through 2018. Results are in ppb.

Figure S1b shows the spatial distribution of the annual mean XCO observed with MOPITTv8. We can observe a latitudinal gradient from north to south with high values over East Asia and central Africa. For the prior, we can see that on average it is higher than the observations, particularly over the tropics (such as Indonesia) but also for northern America such as Alaska in 2015 or Africa. Differences between the prior and the observations over these regions are greater than 34 ppb. The optimized CO concentrations have on average lower difference with the observations, but the concentrations are between 16-20 ppb lower on average across the globe than the MOPITT measurements, except for some regions of Africa, South America in 2015 and East Asia for all years where the concentrations are higher than the MOPITT data. We can particularly observe optimized CO concentrations of 32 ppb greater than the measurements in 2015 over Indonesia and southeast Asia. Looking particularly at Africa and for both tropical south and north Africa, priors XCO are higher than the MOPITT measurements by 32 ppb. The low emissions observed with the posterior compared to the prior could come from this difference between the prior mixing ratio and the measurement and/or the chemistry in TM5 (different from other studies, such as Zheng et al., (2018) which have a multi-species inversion for a more observationally constrained chemistry).

3. And how does the seasonal cycle of these fires look like for Northern and Southern Africa? This can be shown with a figure similar to Figure A4.

The seasonal cycle of Africa is already available in Fig. A1 (now Fig. S3 in supplement information) for the regions: Temperate north Africa, northern tropical Africa, southern tropical Africa, and temperate south Africa. We can see that both prior and posterior emissions are similar for temperate north Africa (with emissions less than 0.12 TgCO/yr), probably caused by the large biases of Sahara dust. The seasonal cycle of the prior and posterior are similar for all regions. For temperate south Africa, we can see higher posterior emissions compared to the prior. For northern tropical Africa, our posterior CO emissions are lower than with the prior for the whole period, even during peaks of emissions, corresponding to annual emissions 25% lower with the posterior than with the prior. While for southern tropical Africa, the maximum peak of emissions is closer each other between the posterior and the prior, particularly in 2016 and 2018, and with a 50-70 TgCO/yr of difference in July-August of 2015 and 2017. The posterior CO emissions are 16% lower than the prior for this region.

4. And finally, is there independent data that support your results? The ‘lack’ of fire emissions in Africa could explain the appearance of the large compensating natural CO2 source in Northern Tropical Africa in your OCO-2 inversions (Figure 11).

We have added, in supplementary information, an evaluation of the CO inversion against TCCON data (Fig. S7, see response to Reviewer #2).

In the evaluation against TCCON, we can observe an underestimation of the posterior CO mixing ratio of ~ -12 ppb in 2015 at the Ascension Island site. However, the a priori CO mixing ratio has an overestimation of 5 ppb in 2015. A similar pattern is found for Reunion Island, with an underestimation of about -7 ppb with the posterior and an overestimation of about 2 ppb with the prior. However, the biases at the Darwin TCCON site give -3 ppb for 2015-2016 (-0.5 ppb for 2017-2018) with the posterior and 20 ppb for 2015-2016 (22 ppb for 2017-2018) with the prior. This gives the impression that our inversion is not getting the best fluxes for Ascension Island, but we can see that this is not the case for other tropical locations. Ascension Island is known to be impacted with Saharan dust and therefore the posterior simulated concentration could be biased due to aerosols.

This underestimation of CO fire with the posterior could explain the larger (smaller) NEE source for MOre emissions compared to other OCO-2 simulations in tropical north and south Africa. For IS inversions, the compensating emissions give a deeper net sink with ISMOre for tropical north Africa compared to other IS simulations. It gives for southern tropical Africa higher net sources compared to IS3re and ISCMS, but lower net sources to IS4re. Now that we have a site-by-site evaluation with TCCON (see Reviewer #2 and Figure 11 in the updated manuscript), we can see that for Ascension Island, ISMOre has the lowest biases with TCCON (-0.55ppm) compared to other posteriors. GFED4re and MOre additionally have the lowest biases with TCCON (-0.80 and -0.83ppm respectively) compared to GFED3re and CMS-GFED3. Even if we cannot conclude which inversion is doing better than the other with the evaluation, it seems that, at least for Ascension Island, ISMOre does better than the other inversions.

5. In addition, I would like to inform the authors of a recent PNAS paper by Ramo et al. (2021) which shows with new 20-m burned area data from Sentinel-2 that the GFED4s emissions for Africa are probably greatly underestimated.

Ramo et al. (2021), https://doi.org/10.1073/pnas.2011160118

We thank the Reviewer for this information. It would be interesting for future work to include these fire emission estimates as a prior and compared the resulting CO emissions with the ones estimated using GFED4.1s. We included this information in the discussion page 36, line 714: A recent study has shown the underestimation for Africa of MODIS burned area and consequently GFED4s, compared to the new Sentinel-2 burned area product (Ramo et al., (2021)). The higher fire posterior emissions observed with previous studies using GFAS as a prior compared to GFED4 (Nechita-Banda et al., (2018)) and the results of Ramo et al., (2021) seems to suggest for future work to carefully choose the CO fire prior used in a CO-CO_2 study. Future work will be done comparing different CO posterior emissions.

6. A similar contradiction arises in Indonesia. In Figure 5a, strangely enough, lower posterior emissions are reported for the different fire types during the intense El Nino fires of 2015. This is the opposite of what Nechita-Banda et al. (2018) found in their MOPITT CO inversion study. They found emissions up to 120 Tg CO, about 1.5 times greater than what GFED4s provided. This difference should also be addressed and explained in your paper.

The reviewer compared our MOPITT CO inversion against the inversions done in the paper of Nechita-Banda et al., 2018 mentioning that our results corresponding to the 2015 annual emissions are different to the estimation of Nechita-Banda et al., 2018. However, it is important to remark that the temporal scale, spatial scale, data assimilated and prior used are different between the two studies. First, while our study looks at annual emissions, Nechita-Banda et al., 2018 looks for the brief period of mid-August to mid-November 2015. Their study does not include the whole year 2015. Second, the regions of interest over Southeast Asia are different between our two studies. Their studies focused on the Indonesia and Central tropical Asia, while our study has two regions that divide Indonesia and Central Asia (please refer to Fig. 4 of our manuscript). Finally, the data used and assimilated in both studies are not the same. In our paper, for the CO inversion, we only assimilated MOPITT CO TIR-NIR version 8 data. However, in the study of Nechita-Banda et al., 2018, they used and assimilated both MOPITT CO NIR-TIR version 7 data and the ground-based NOAA data. We can read in their paper, page 4, section (d): "*The main results described in this paper are from the 'IASI' and 'MOPITT' simulations, where IASI and, respectively, MOPITT level 2 data are used together with NOAA surface stations, to optimize CO emissions*". NOAA ground-based data have been shown to have higher CO biomass burning emissions compared to the prior when assimilated at global scale (Hooghiemstra et al., 2012). The use of these ground base data in Nechita-Banda et al., 2018, together with MOPITT data could explain the higher emissions than with MOPITT alone. Additionally, it is important to notify that our posterior

MOPITT emissions are done using the prior GFED4.1s, while in Nechita-Banda et al., 2018 they used and mentioned "*Most simulations presented use GFAS v.1.3 as prior biomass burning emissions, except for 'IASI OH GFED', which uses GFED4s as prior*". Additionally, they have adjusted the GFED peat emissions with a higher CO/NOx ratio than typically used in GFED. But, in their Fig. 4, we can see for the total Indonesia region that the posterior emissions IASI using GFAS (called IASI OH) has higher emissions than the posterior emissions using GFED as prior (called IASI OH GFED). Additionally, in their Fig.5, we can see, for their period of study, higher CO fire emissions with prior GFAS than with prior GFED. Their posterior CO fire emissions (opt MOPITT and opt IASI) have similar seasonality than prior GFAS and consequently give higher emissions than the prior GFED. In their study, they do not present results of posterior emissions using only MOPITT data with GFED prior. Using both MOPITT-NOAA data and the prior GFAS in their inversions, could have given higher emissions than using a MOPITT-GFED inversion. We cannot, consequently, compared our CO fire estimate with Nechita-Banda et al., 2018.

In their paper, we can find a comparison with two other studies, Yin et al., 2016 and Huijnen et al., 2015, which both studies focused on similar region and assimilated MOPITT data using the prior GFAS in their inversion. In Huijnen et al., 2015 paper, we can see in Fig.4 that their posterior emissions are lower for the whole period of study than their prior (GFAS), contradiction to what Nechita-Banda et al., 2018 found. When Nechita-Banda et al., 2018 extrapolated the results of Huijnen et al, 2015 to the same period they used, they found 96 TgCO (17 TgCO lower).

Finally, it is important to mention that no evaluation against independent data are performed in their study. Therefore, there is no result proving that their results are more accurate or more trustworthy than our results.

7. Besides that, it also contradicts with what is shown in the lowest panel of Figure 9 of your paper. The blue (GFED4re) and red (MOoptre) bars are (almost) identical for 2015 (and the other 3 years). And finally, this result contradicts with the statement at line 694: "the fire estimated from MOPITT CO emissions are stronger than with GFED4 emissions". Again, according to Figure 5a, MOPITT CO emissions are actually smaller.

In Figure 5 (now Figure 4), we represented the biomass burning of MOPITT CO posterior and GFED4.1s prior emissions partitioned by vegetation types. The partition of the vegetation types is done using the fraction of dry matter available in the GFED product. From the MOPITT CO Posterior emissions, we calculated and estimated the average CO2 fire emissions (corresponding to FIREMo in our study). However, in some CO2 inversion, biofuel is included with the biomass burning emissions. Biofuel refers here as the CASA-GFED3 (Ott, 2020) product of the anthropogenic burning of harvested wood (van der Werf et al., 2010). This product is calculated as the population density times national per capita fuel consumption estimates while being constrained by the total available coarse woody debris at each model time step. Consequently, we summed each fire and biofuel emissions to get FIRE3, FIRE4 and FIREMO used in our CO2 inversions. We have already included a sentence in the manuscript "*All CO2 FIRE priors include both biomass and biofuel burning*" referring to this inclusion. But since this seems to not be informative, we also added this information in Table 3, with corresponding FIRE emissions. It is important then to remember that biofuel emissions are used in our CO2 results for the FIRE emissions.

Regarding line 694, our sentence: "*the fire estimated from MOPITT CO emissions are stronger than with GFED4 emissions but lower than GFED3 emissions*" was referencing to CO2 fire estimated of Figures 9 and A3, the only section where we compared all together fire estimates from MOPITT, GFED4 and GFED3. We are sorry for the confusion, we should have mentioned this sentence by "In 2015, during the onset of the El Nino event which caused intense fires over Indonesia, FIREMo are stronger than with FIRE4 emissions but lower than FIRE3 emissions." In the comments below, we reply and calculate the conversion from CO to CO2 for FIREMo, which is consistent with our results. We remind that we did only the conversion for the FIREMo. We did not convert our CO prior emissions into CO2 fire emissions but used the GFED4.1s and GFED3 products available from the CASA-GFED products.

8. Given these two major inconsistencies, I wonder how much we can trust the results of the fire inversions for the other regions. For example, the strong reduction in fire emissions in the boreal regions is also striking and is somewhat inconsistent with the record-breaking fires we have seen in recent years in these regions.

For the boreal regions, we indeed have lower posterior CO fire emissions than the prior. But this seems to come from the measurements. You can see Fig. S1b, that the MOPITT measurements do not see any signal from the boreal regions. The comparison PRIOR-MOPITT in this figure shows that the prior mixing ratio has higher concentrations than the measurements over Alaska and the boreal regions.

There is then an underestimation of the measurement compared to the prior for this region. This underestimation seems to emerge in the posterior mixing ratio and could explain the lower emissions with the posterior observed in the fire CO emissions. The evaluation with TCCON data (Figure S7) shows indeed an underestimation globally of 7ppb with the posterior mixing ratio, but the prior has a larger bias (of around 13ppb).

9. I also noticed that the breakdown of fire emissions in Figure 5 does not agree at all with the reported CO emissions by GFED4s. As an example, if I look at Southern Tropical Asia in panel 5a, the breakdown between GFED4s agricultural, deforestation, savanna, and peat fires is roughly: 5, 38, 25 and 22 Tg CO/yr, respectively. However, if I look at the published GFED4s emission tables for 2015 for the EQAS region (roughly similar to your STA region) I get a completely different set of emissions per fire type: 0.7, 31, 3.8 and 74.2 Tg CO/yr, respectively. So in words, GFED4s reports a much larger source of CO from peat fires and a smaller source of CO from savanna/grasslands than you. Therefore, please carefully check Figure 5 for possible errors for all the regions.

GFED4s tables: https://www.geo.vu.nl/~gwerf/GFED/GFED4/tables/GFED4.1s_CO.txt

To verify that vegetation partition was correctly applied to the CO prior, Figure R1 below shows the difference between the prior CO emissions for January and June 2015 on a global scale minus the sum of the emissions partitioned by vegetation.

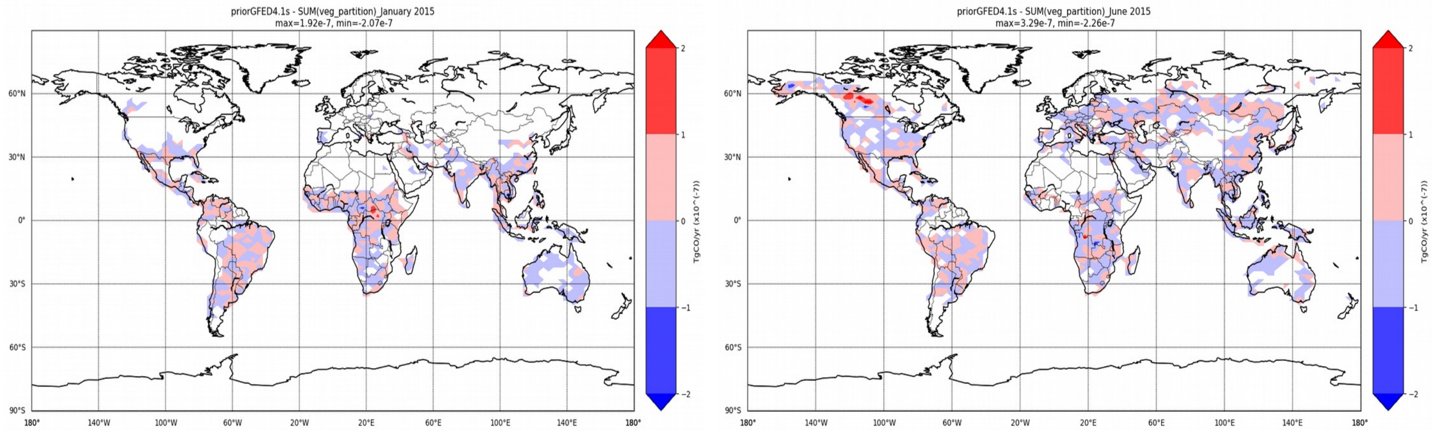

Figure R1. Difference between the prior CO Fire emissions (GFED4.1s) versus the sum of the emissions partitioned by vegetation for January 2015 (left) and June 2015 (right). Values are in TgCO/yr ($x10^{-7}$).

We can see on the figure that the differences are very small (in the order of 3x10-7 TgCO/yr) and so quasi in-existent. The vegetation partitioned shown in our Figure 5 from the preprint is then correct.

The difference between the values reported on the GFED website and our prior could be caused by the difference in spatial scales. The GFED reported on the website are reported at the MODIS resolution and then aggregated by region, while our estimates are reported at 3x2 spatial resolution.

10. Similar issues also apply to Fig. A3. The contributions from the different fire types do not agree with the reported values in Fig. 5.

We need to remind the readers that we calculated and converted the CO_2 fire emissions from CO fire emissions only for our FIREMO estimates. We did not convert the emissions for the FIRE4 (for GFED4.1s) and FIRE3 (for GFED3). The FIRE4 and FIRE3 have been taken directly from the CASA-GFED emissions inventories. Additionally, like previously mentioned, this is potentially a difference in spatial scales. The GFED emissions are calculated at the resolution of MODIS pixels, while the FIREMO ones are done at 3x2 degrees resolution. Then, like already mentioned in the manuscript, all CO_2 inversions are done at 6x4 degrees resolution.

We double checked if the FIREMO reported for the CO_2 results are correctly calculated from the CO fire emissions, which indeed they are. So for instance, if we take the CO fire emissions of our CO posterior emissions (MOPITT) reported in Figure 5 (now Figure 4) over southern tropical Asia in 2015, which is approximately 15 TgCO/yr for the peat lands, we need to apply the following equation (as already explained in our manuscript in section 2.3.1.b):

$$\frac{\frac{12}{28} * CO(TgCO/yr) * 10^{-3}}{ER_{peat}} = CO_2(PgC/yr) \rightarrow \quad \frac{\frac{12}{28} * 15 * 10^{-3}}{0.194} = 0.033\, PgC/yr$$ which corresponds approximately to our figure A3.

If we do the same with the CO values of deforestation, agriculture and savanna which are approximately 26 TgCO/yr, 3 TgCO/yr and 23 TgCO/yr, we get respectively 0.13, 0.013 and 0.17 PgC/yr. From these values, the biofuel emissions need to be added to have exactly the values shown in Figure A3 (now Figure S5). These values are practically similar to our estimate in Figure A3 (now Figure S5).

11. To better understand the reported posterior CO fire emissions it would also be helpful if the TCCON comparisons for CO are provided to the reader (in a supplementary document).

We have already answered this comment with our reply to Referee #2. We added in the supplementary document, the TCCON evaluation of our prior and posterior CO mixing ratio. We ask the reader to refer to the comments of Referee #2 for further information.

12. Finally, I would like to reiterate Meinrat Andreae's comments that the emission factors should indeed be replaced by the updated ones reported in Andreae (2019).

We answered to this comment with our reply to Meinrat Andreae’s comment. We ask the reader to refer to Andreae’s comment for further information.

2. Role of fire emissions on the optimized NEE

13. How can we benefit from this more tedious 2-step inversion approach? The answer is somewhat lost in the lengthy description of the results. If I simply look at Figures 11, 12 and 13, the differences in NEE between the different inversion experiments (with the same data constraints) are small. The largest differences originate from using either OCO-2 or in situ CO2 data for reasons discussed in the paper (like differences in data coverage). However, with either data constraint, the optimized CO2 fire emissions seem to have a small impact on the NEE in comparison to the unoptimized CO2 fire emissions.

As described in the abstract, both MOPITT derived fires and GFED4s fires provide larger net sources in the tropics.What specifically did we learn from the MOPITT experiment? Also the comparison with independent TCCON data in Table 6 does not show significant differences in comparison to the other inversions if you use the optimized fire emissions in a NEE inversion. A more in depth discussion about the wider application of this methodology is necessary and should be mentioned in the abstract.

We thank the reviewer for this comment. We indeed shortened the manuscript as advised in the following comments and by Reviewer#2. However, we added in the evaluation section a comparison between the priors and posteriors with the OCO-2 retrievals and IS measurements to see how the inversions fit the data. We also changed the evaluation with TCCON to have a site-by-site evaluation. Regarding the abstract, we changed the sentences page 1 line 20:

"Evaluation with TCCON suggests that the re-balanced posterior simulated give biases and accuracy very close each other where biases have decreased and variability matches better the validation data than with the CASA-GFED3. Further work is needed to improve prior fluxes in Tropical regions where fires are a significant component."

as following:

"Evaluation with TCCON data shows lower biases with the three re-balanced priors than with the prior using CASA-GFED3. However, posteriors have accuracies very close each other, making difficult the conclusion of which simulation is better than the other. One major conclusion from this work is the strong constrain at global scale of the data assimilated compared to the fire prior used. But results in the tropical regions suggest sensitivity to the fire prior for both the IS and OCO-2 inversions. Further work is needed to improve prior fluxes in tropical regions where fires are a significant component. Finally, even if the inversions using the FIREMO prior did enhance the biases over some TCCON sites, it is not the case for the globe. This study consequently push forward the development of a CO-CO2 joint inversion with multi-observations for possible stronger constraint in posterior CO2 fire and biospheric emissions. ”

3. Readability of the paper

14. The paper would benefit from some additional editing. Mainly to correct and shorten the long (and sometimes awkward) sentences and to trim down some of the lengthy descriptions. In particular, the Results section should focus more on the main results of the paper and be presented in a more concise and logical way. Now I find it difficult to quickly get the gist of the paper in the whole list of comparisons between the inversions and between the different data constraints used. The fact that it is a two-step inversion makes that even more difficult. The posterior of the first inversion becomes the prior of the second inversion, and that blurs the distinction between ‘prior’ and ‘posterior’ labels. There is so much work in this paper that I wonder if it wouldn’t be better to split the paper into two separate papers: one for the estimation of fire emissions with MOPITT CO constraints, and a follow-up study on the impact on NEE with OCO-2 and in situ data, but that's something up the authors to decide. You can perhaps choose to move some the detailed comparisons to a supplementary document where the reader can find additional information about the main results.

Thank you for this remarks. We decided to not separate the paper as the comparison among the different simulations was to see any impact or not of using CO information in CO2 inversions with two different set of assimilated data. We consequently chose to move some of information in a supplementary document to make the paper easier to read. For more clarity also, we have added some diagrams in the methodology section.

15. I also noticed at lot of inconsistencies between labels reported in the caption, in the figures and elsewhere in the tables and the main text. For example, I came across multiple names for the same inversion type: CMS-GFED3 (in Table 5), CMS (e.g. in Fig. 12) and OCOcms (in Fig. 5a). Another example: MOPITTopt (e.g. line 455), MOoptre (e.g. Fig. A4), and MOre (e.g. Table 4). I find this distracting and should be fixed in the revised manuscript.

Sometimes you interchangeably use the inversion name (e.g. MOre) when you refer to a flux and vice versa. For example at line 454 you write: “The prior categories shown are fire, NEE and net fluxes for the prior GFED4, GFED3, MOPITTopt and CMS-GFED3”.

If you follow your own Table 5, this should change to: “The prior categories shown are fire and NEE for GFED4re, GFED3re, MOre and CMS-GFED3”.

We corrected all of the annotation in the manuscript and particularly named the priors, and the posteriors.

16. I also believe the number of figures should be reduced (currently 14 excluding the Appendix figures). You can merge a number of figures into a single figure. For example, Figure 8 and 9 should be combined into 5 squared panels. That provides an overview of all emission estimates on a single page. Similarly, Figures 11, 12 and 13 can be combined into a single page filling figure with 8 panels.

We reduced the number of figures and instead of splitting the figures by region, we only use the figure that was containing all regions. This figure was previously available in the annex.

17. Finally, shortening of the region labels with No and So in the figure titles is often not necessary, is confusing, and it doesn’t look very pretty. If there is enough space write the labels in full. For example, So Trop So America should become Southern Tropical South America.

We corrected the labels for all figures.

18. In the specific comments below I give more examples of long sentences, wrong usage of inversion names, and other specific errors.

We thank the reviewer for these specific comments.

Specific comments:

Line 6. This part is not clear. In the end are the CO_2 NEE and ocean fluxes optimized with OCO-2 and insitu data or only NEE? This should be stated very clearly in the abstract.

Suggestion:

These optimized CO_2 fire emissions (FIREMo) are used to re-balance the Net Ecosystem Exchange (NEEmo) and respiration (Rmo) with the global CO_2 growth rate. Subsequently, in a second step, these rebalanced fluxes are used as priors for an inversion to derive the NEE and ocean fluxes constrained either by the Orbiting Carbon Observatory 2 (OCO-2) v9 or by in situ CO_2 data.

The NEE and ocean fluxes are optimized with OCO-2 and IS. This was mentioned page 15, line 357. We took the suggestion in consideration and changed the sentence with: These optimized CO_2 fire emissions (FIREMo) are used to re-balance the CO_2 Net Ecosystem Exchange (NEEmo) and respiration (Rmo) with the global CO_2 growth rate. Subsequently, in a second step, these rebalanced fluxes are used as priors for a CO_2 inversion to derive the NEE and ocean fluxes constrained either by the Orbiting Carbon Observatory 2 (OCO-2) v9 or by in situ CO_2 data.

Line 11. Be consistent throughout the paper with labels. Use either CASA-GFED3 or GFED3. We corrected this through the all paper.

Line 12. “Results show…” Unclear what this sentence is trying to say. Does “Results” refer here to the evaluation with TCCON? Or does it refer to the flux estimates?

The way I read it is that the posterior flux estimates (whether you mean NEE or fire is also unclear) are more robust (i.e. similar) than the different prior flux estimates. For clarity, we changed the sentence to: Comparison of the flux estimates show that at global scale posterior net flux estimates are more robust than the different prior flux estimates. However, at regional scale, we can observe differences in fire emissions among the priors, resulting in large adjustments in the Net Ecosystem Exchange (NEE) to match the fires and observations.

Lines 16-20. I find this short recap of the main results quite hard to read because so many geographical regions and elements of 2-step inversion approach are compressed in 2 sentences (GFED4s, MOPITT CO, OCO-2, insitu data). Please rewrite.

A suggestion:

Slightly larger net CO2 sources are derived with posterior fire emissions in the OCO-2 inversion, in particular for most Tropical regions during 2015 El Nino year. Similarly, larger net CO2 sources are also derived with posterior fire emissions in the in-situ data inversion for Tropical Asia.

We rewrite this lines with: Slightly larger net CO2 sources are derived with posterior fire emissions using either FIRE4 or FIREMo in the OCO-2 inversion, in particular for most Tropical regions during 2015 El Nino year. Similarly, larger net CO2 sources are also derived with posterior fire emissions in the in-situ data inversion for Tropical Asia.

Line 21. Use either: ‘re-balanced posterior simulation’ or ‘re-balanced posterior simulated concentrations’ . We changed this sentence regarding a comment from Reviewer #2: Evaluation with TCCON data shows lower biases with the three re-balanced priors than with the priorCMS. However, posteriors simulated concentrations give biases and accuracy very close to each other, making difficult a conclusion of which simulation is better than the other.

Line 21. ‘very close **to** each other’ Done

Line 34. Be aware that since 2017 GFED4s emissions are not based anymore on direct burned area datasets, but instead based on relationships between MODIS active fire detections and GFED4s emissions for the period 2003-2016. This is because the underlying burned area dataset has been upgraded in the meantime from Collection 5.1 to Collection 6, making it incompatible for usage in GFED4s. GFED4s emissions from 2017 onward are therefore called GFED4s-beta emissions.

See: https://www.geo.vu.nl/~gwerf/GFED/GFED4/Readme.pdf

We modified the sentence “The first uses total fuel consumption per product of the burned area and the fuel consumption per unit area deduced from the burned area and active fires products of the Moderate Resolution Imaging Spectroradiometer (MODIS).” by:

"The first uses, since 2017, total fuel consumption per product of the burned area and the fuel consumption per unit area deduced from the burned area and active fires products of the Moderate Resolution Imaging Spectroradiometer (MODIS). Previous years were based directly on burned area datasets."

Line 51. Induce Done

Line 81. There are many examples you write double plural "emissions sources", "emissions estimates", "emissions inventories". Although I'm not an English native speaker, I think it's grammatically better if you write it as "emission sources" or "emission estimates". Corrected

Line 98. "and the post-event" I suggest "and the subsequent years" Changed

Line 106. "The importance of these results for conclusions …". I think this is grammatically incorrect. Please rewrite. Something like: "The importance of these inversion results are discussed in Section 4." Sentence changed to: The importance of these inversion results are discussed in Section 4.

Line 110. Please provide a clearer breakdown of the 2-step approach. A suggestion:

Our inversions are performed in sequence: (1) we assimilate total column CO retrievals from the MOPITT v8 products to produce optimized CO fluxes, which are used to update the assumed CO2 fire emissions, and then (2) we assimilate either total column CO2 from OCO-2 version 9 retrievals or CO2 in situ data to produce optimized CO2 NEE fluxes.

We considered this suggestion and wrote: Our inversions are performed in sequence: (1) we assimilate total column CO retrievals from the MOPITT v8 products to produce optimized CO fluxes, which are used to update the assumed CO2 fire emissions, and then (2) we assimilate either total column CO2 from OCO-2 version 9 retrievals or CO2 in situ data to produce optimized CO2 NEE and ocean fluxes.

Line 113. validation (singular) Corrected

Line 124. "…allowing a well-understood of its continuity and consistency ". Please rewrite. We modified the sentence with: However, MOPITT products have been consistently validated against airborne vertical profiles and ground based measurements, allowing a well-understood product

Line 177. Awkward sentence. Perhaps start it like this: "Despite the known shortcomings (biases) of satellite data, several studies have preferred to use satellite data over the Tropics to take full advantage of the improved spatial coverage." Changed

Line 182. at smaller spatial and temporal scales Changed

Line 214. suggestion: "…shows the site locations over the globe." Changed

Line 226. Change to "the global in situ network" Added

Line 233. …the corresponding satellite and in situ data. Added

Line 236. R covariance structure is not discussed in the paper. We have added the sentence page 12 line 290: The errors are assumed uncorrelated leading to a diagonal observational error covariance matrix **R**.

Line 262-264. These 2 sentences should be merged to the same paragraph. One of the reviewer's comments concerned the length of the manuscript and recommended reducing it. To shorten the manuscript, we have therefore moved lines 264 to 277 to the appendix.

Line 272. Change to "fire carbon emissions with 11%" Changed

Line 293-296. This part needs some editing. Also, I'm not sure what you trying to say here. Do you mean perhaps if an optimized CO flux for a pixel becomes twice as large (after inversion with MOPITT CO) you scale up the fluxes of the underlying vegetation types with a factor of 2? This section needs indeed some editing and clarification. For this purpose, we added in the manuscript a flowchart (Fig. 1) about the CO inversion and the vegetation partitioning used for the FIREMo calculation.

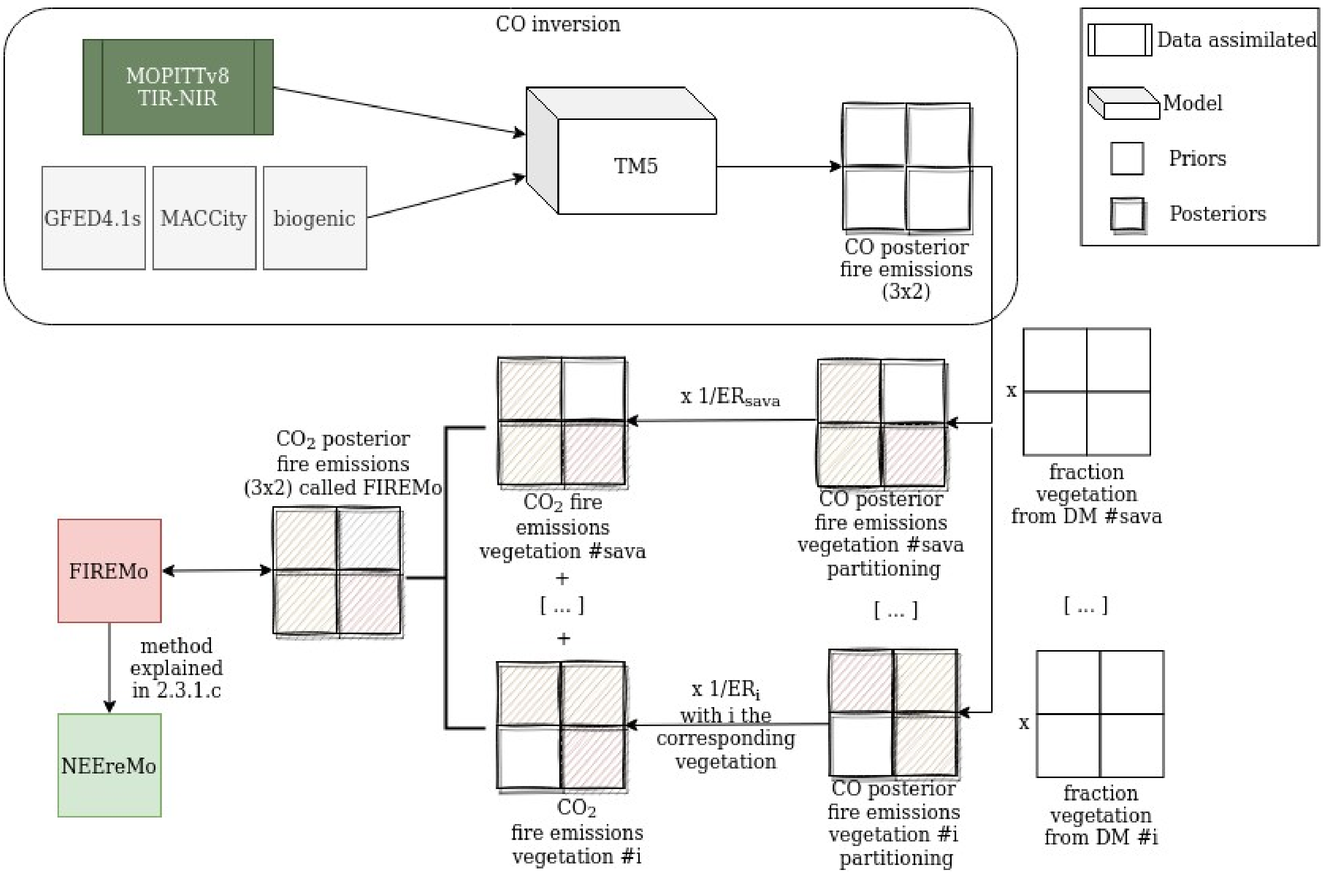

Figure 1. Flowchart of the FIREMo calculation

We also changed the paragraph to:

In this section, we describe the computation of our optimized prior fire emission (FIREMo) which we will use to observe the impact of CO fire emissions in posterior CO2 biospheric fluxes. The steps of the FIREMo calculation are shown in Fig. 1. For each pixel of CO posterior fire emissions, we applied a vegetation fraction based on the dry matter product (DM) of GFED4.1s. We obtained fire emissions for each monthly vegetation type (savanna, boreal forests, peat, temperate forests, deforestation and agriculture waste). Figure S1 shows GFED DM vegetation type for each year over land, where each pixel represents one or more vegetation types.

Line 296-297. Suggestion: "Figure 3 shows for instance the GFED vegetation type for each year, where each pixel represents one or several vegetation types." We modified the sentence by: Figure S1 shows GFED DM vegetation type for each year over land, where each pixel represents one or more vegetation types.

Line 299. As mentioned Andreae. Use the new updated EFs published in Andreae (2019). We already replied to this comment in Andreae's comment. The CASA-GFED3 product does not use the new updated EF published in Andreae (2019) and consequently, for consistency among the fire product used, we did not use the new EF product in our FIREMo product.

Line 309. "type per grid box". This is at 3x2 resolution? If so, state that explicitly. Changed

Line 311. "in balance with fire estimate". You mean in balance with the atmospheric CO_2 trend, or not? We changed the sentence by: We used this FIREMo as a fire prior emissions in CO_2 inversions along with a re-balanced respiration and NEE (in balance with fire estimate), using the parameterization described in the following section 2.3.1.c.

Line 318. Unclear. Are the ocean fluxes also optimized or not? We modified the sentence by: Ocean fluxes are taken from Takahashi et al. (2009). They are assumed to have an uncertainty variance of 50%. Both biospheric and oceanic emissions are optimized in the CO_2 inversions.

Line 323. You mean CASA-GFED3 We meant GEOS-Carb CASA-GFED3 project.

Line 326. "gross ecosystem exchange". Is this the correct terminology? Should it not be the Net Primary Production (NPP)? NPP is equal to the sum of gross primary production (GPP) and autotrophic (maintenance) respiration (Ra). See below.

$NEE = GPP+Rh+Ra=Rh+NPP$

NPP = GPP+Ra

So, we can read in our paper that the net ecosystem exchange (NEE) is expressed as the sum of heterotrophic respiration (Rh) and gross ecosystem exchange (GEE) : NEE= Rh + GEE

Which corresponds with what you wrote : NEE= GPP + Ra + Rh = GEE + Rh. Our equation is right, and the Rh is the respiration we have balanced with the fire, approach used in GEOS-Carb CASA-GFED3 (Ott , 2020).

eq. 2. What is the meaning of number 3? Because it is based on CASA-GFED3? Yes it is referencing to CASA-GFED3. The text above the equation already mentioned the meaning of elements used in eq.2. We modified Geos-Carb into CASA-GFED3.

Line 328. What is the meaning of MOPITTOpt and RMo? First time these parameters are mentioned. We corrected that with FIREMo and RhMo (respiration linked to FIREMo)

Table 3: Something is off with the calculated values of FIRE4 (GFED4s). When I calculate the emissions myself from the official GFED4s tables I calculate for 2017, 2018 and 2019 very different emissions than what is reported in Table 3 (highlighted here in bold). I take the global emissions from https://www.geo.vu.nl/~gwerf/GFED/GFED4/tables/GFED4.1s_CO2.txt and subsequently convert them to carbon emissions with 12/44 ratio.

FIRE4 1.88 2.09 1.73 **1.78 1.69 2.13**

We corrected the values as indeed the values were not correctly written. We corrected both GFED4 and FIRE4.

It would also be helpful to see the estimates of NEEre3, NEEre4 and NEEreMO, and CMS-Carb in Table 3.

We added the estimates in Table 3. We also decided to remove 2014 and 2019 values are the full years were not used in the inversion (spin up period).

There is also another issue with FIRE4 and FIRE3 estimates in Table 3. They do not match with the blue and green bars in Figure 7. All estimates in Fig. 7a exceed 2 PgC/yr, which is not the case in Table 3.

Indeed, as we mentioned page 15 line 351 “All CO2 FIRE priors include both biomass and biofuel burning. “ The biofuel of table 3 is additionned to the fire emissions to get the FIRE emissions used in the inversions. When we addition the fire and biofuels emissions of table 3, we find the values shown in Figure 7. We changed the table3 by adding GFED3, GFED4, FIRE3 and FIRE4 values.

Line 380-423. Very difficult to read. A lot of inconsistencies in region labels, and redundant information in the text. I suggest to keep it simple and concise by briefly discussing the results in separate short paragraphs for each continent. Just keep it simple. Highlight that the optimized fires are generally smaller than prior fires.

This section also omits important large differences between GFED4s and the posterior fires (see my earlier concerns). Especially the differences for the boreal regions, Africa, and Indonesia are striking and inconsistent with the recent literature. This should be addressed.

As discussed in the comment earlier, the studies of Zheng et al., (2018) and Nechita-Banda et al., (2019) did not use the same data in their inversions or assimilated additional data. Zheng et al., (2019) shows that their inversions set-up gives higher emissions than other inversions. Moreover, no evaluation is available to determine which results are the most reliable. We just added a few sentences in our manuscript about the differences between our study and their results. We rewrote the paragraph as:

Figure 5 shows the annual CO posterior and prior fire emissions split by vegetation combustion across the globe and by OCO-2 MIP regions. Overall, it can be seen that depending on the region, the assimilation of MOPITT data involves less or more CO emissions compared to the prior GFED4.1s.

For North Temperate America, posterior emissions remain close to the prior estimates, suggesting that the inferred emissions are consistent with GFED4.1s. Comparable results are also observed for Temperate North Africa. However, this region is known to have a lot of Saharan dust transported across the Atlantic Ocean and towards Europe most of the year, which could explain the posterior emissions being close to the prior as those MOPITT soundings have largely been removed by pre-screeners. North Tropical Africa is not only affected by dust, but it is also largely affected by clouds during the wet season of the African monsoon (from May to August), which could lead to errors in retrievals that pass the pre-screeners. The combination of clouds and dust could explain the MOPITT posterior fires having lower emissions than the prior GFED4.1s estimate. But further investigation into North Tropical Africa is needed. Even though the prior is higher than the posterior for tropical Africa, in opposition to the previous multi-species study of Zheng et al., (2018), the posterior emissions better fit MOPITT measurement than the prior (Fig. S4). Tropical south America (including North Tropical South America and South Tropical South America) is also known to have cloud coverage limiting satellite observations. We however observe similar emissions between the prior and the posterior for the northern region, with slightly higher emissions for MOPITT. For the southern region, differences between the prior and the posterior are strong. The cloud coverage might explain this behavior, but further investigation are needed for these two regions.

The discrepancies observed for Eurasia temperate between MOPITT and GFED4.1s could be that the vegetation type is not well represented for these regions. As mentioned in Pechony et al., (2013), agriculture and savanna vegetation types might not be the dominant burning vegetation type over North Africa and the Middle East, since these regions have seen an increase in croplands area well control by human activities and so burn rarely. However, Kazakhstan is a region of temperate forest often dominated by fires (Venevsky et al., (2019)), a characteristic that MOPITT posterior emissions seem to observe as mush as the prior GFED4.1s.

We can also observe that over Northern Tropical Asia, MOPITT fire emissions are higher than GFED4.1s (see Fig. 4 and Fig. S7). This is observed for all years, where MOPITT emissions are almost 5 TgCO/yr (2 TgCO/yr) for savanna (for the other vegetation types) higher than from GFED4.1s. As mentioned in Petron et al., (2002) and Arellano et al., (2004), CO emissions in Northern Tropical Asia are significantly underestimated in current inventories. Previous studies have shown that peat area and depth, producing large amount of carbon (~0.60 PgC/yr which represents 26% of the total carbon fire emissions, (Nechita-Banda et al., (2018)), were found to have significant uncertainties in Indonesia in the emissions inventories (Lohberger et al., 2017, Hooijer et al., 2013). Our posterior have lower emissions than the prior for southern tropical Asia, in contradiction to what Nechit-Banda et al., (2019) observed. However, Nechita-Banda et al., (2019) assimilated MOPITT and NOAA observations and used GFAS as fire priors, an inversions set-up different to what we used. Additionally, no evaluation against independent data have been performed in their study to determine if their results are more trustworthy than our results.

Our posterior can capture the seasonality of peat fires over Indonesia in comparison to GFED4.1s. Figure S6 shows for Southern Tropical Asia (mainly visible in 2015 due to the large emissions) that GFED4.1s have a fire peak earlier than MOPITT. VanderLaan-Luijkx et al., (2015) and Nechita-banda et al., (2018) hypothesized that GFED4.1s might not capture the timing of emissions over area with peat fires due to the use of burned area, which may be more sensitive to the initial stages of the fire than to the continued burning.

Figure 5. - Put labels (a) 2015, (b) 2016, (c) 2017 and (d) 2018 on top of each panel

Typo in legend: moppit should be MOPITT

In the legend make MOPITT and GFED4s transparant white, not grey (as this color is already used for agri).

I suggest to include a bar graph in each panel that shows the global emissions for the different land types.

We modified the figure, following comments from Reviewer #2, with only a bar plot figures. We modified the legend as well as suggesting here. The figure can be find with the comments of Reviewer #2. A figure with the global emissions for each regions was already present in the manuscript in Annex. We moved this figure in supplementary information (Fig. S???).

Line 428. You mean "The first one" Changed

Line 438. temperate without capital Changed

Figure 7: Please put the experiment labels in the middle of each grouping of bars. In addition, label the 4 panels with a, b, c and d. Similar comments apply to the other bar plots. We changed the label in all bar plots. However, regarding the space available in each panel. We did not modified the legend.

Section 3.2.1. Overall I believe this section can be trimmed down and made more concise. Done

Figure 8 and 9 should be combined into 5 squared panels. Done

Line 474. "global fires emissions" => "global fire emissions" Done

"We can also observed" => "We can also observe" Done

Line 477. "FIREMo observes less emissions" => e.g. "FIREMo yields less emissions" Done

Line 485. Example of a somewhat tedious number of sentences that should be condensed to one sentence.

"The larger emissions with FIREMo compared to FIRE4 over tropical Asia comes mainly from some specific vegetation. The main vegetation type in this region is savanna and we can observe that for the CO2 prior emissions, FIREMo has the higher flux for Northern tropical Asia (Southern tropical Asia) compared to FIRE3 and FIRE4 (FIRE4 respectively) for savanna but also for agriculture and deforestation (see Fig. A3)."

Besides the sentence structure, I disagree with what it says. Figure 9c shows FIREmo emissions are of similar magnitude as FIRE3 and FIRE4. We modified the paragraph. However, Fig A3 shows higher savanna emissions with FIREMO than with FIRE4. We then kept this sentence and moved Fig. A3 in supplementary information:

The larger emissions with FIREMo compared to FIRE4 over tropical Asia comes mainly from savanna (the main vegetation type in this region, see Fig. S???).

Line 493. capture => captured Done

Line 495 and several other lines. "the GFED" => "GFED" Done

Line 498. Smokes => smoke Done

Line 510. "…between OCO-2 and IS inversions the larger ensemble…" => "…between OCO-2 and IS inversions detailed in Crowell et al. (2019) and Peiro et al. (2022)" Done

Line 517. "…for the Tropics with larger sources for the OCO-2 inversions…". Is this not simply because FIREmo is much smaller in tropics than FIRE4, and thus we see a compensating effect by increasing the NEE source? This sentence does not refer to a specific simulation but compares the OCO-2 inversions with the IS inversions. The sentence was specifically "The net sinks observed with the in situ inversions are weaker than OCO-2 for 2016, 2017, and 2018, and the year-to-year variations are significantly larger than the OCO-2 results. Similar behavior is observed for the SH Ext and opposite behavior for the Tropics with larger sources for the OCO-2 inversions, which could be related to cross-talk between the zonal bands given the sparse coverage of in situ data in the Tropics. " Fires are the same between the IS and OCO-2 inversions. So the difference in sink or source between the IS and OCO-2 inversions are mainly driven by the assimilated data. As we mentioned, the differences in behavior could be related to the sparse coverage of in situ data in the tropics.

Line 517. "…opposite behavior for the Tropics" This part should be the beginning of a new sentence. We changed by: Similar behavior is observed for the SH Ext. However, opposite behavior is observed for the Tropics, with larger sources for the OCO-2 inversions. The differences between both inversions could be related to cross-talk between the zonal bands given the sparse coverage of in situ data in the Tropics.

Line 525. "we can again observe a consistency in OCO-2 across the priors "

You do perhaps mean " consistency between OCO-2 and the priors" ? We did mean across the priors. Similarly to line 520. We changed the OCO-2 part of the sentence however by: We can again observe a consistency across the priors of the OCO-2 inversions.

Line 526. "(with sources for OCO-2 inversion)". I don't understand what this means. We reformulated this sentence: MOre and ISMOre have a smaller sink in 2015 (with sources for OCO-2 inversion) compared to the other inversions in order to balance the 0.5 PgC/yr smaller fires that FIREMo gives.

By:

More has a smaller sinks in 2017 and 2018, but has a source in 2015 (larger in 2016) compared to the other inversions, in order to balance the 0.5 PgC/yr smaller fires that FIREMo gives.

Line 530. "This could then explain why we observe stronger sinks with in situ than OCO-2 posterior NEE emissions. " Which figure can we observe this? We can observe this in Figure 10 of the pre-print (now Fig. ???.c).

Line 532. "different with a Tropical sink in all years except in 2015 and 2016 ". Where can I see this? In figure 10? Indeed, in Figure 10.c, we can observe that the net fluxes of IS give a net sink in 2017 and 2018, but a net sources for 2015 and 2016 (with the exception of ISCMS which has a net sink very close to 0).

Line 522-544. This part needs to be rewritten. I feel that the key message is somewhat lost in this extensive summary of differences between inversions and regions. Make sure the following items are expressed in a concise manner:

- Sinks of OCOcms and IScms are generally weaker than the other inversions. Suggest sensitivity from the imposed AGR.

- Global sinks are larger with in situ data, which is largely driven by larger sinks in the tropics. Possible culprit: sparse data coverage.

- While OCO2 inversions show larger sources over the tropics and larger sinks over SH Ext. Could be a compensating effect for the scaled down fire emissions in the tropics. Done

Line 531. At least from Figure 10 I conclude that the in situ inversions yield for all 4 years a tropical sink, not only 2017 and 2018 as you write. I also see sinks up to 2.5 PgC/yr. How does this relate to the reported sink of -0.5 PgC/yr? This sentence was referring to the net fluxes (see line 532), which indeed the net sinks of IS inversions are around -0.5 PgC/yr. But, as expressed to the previous comment, we rewrote this paragraph.

Line 548. "…we can see that the OCO-2 inversions have deeper net sinks over the Boreal regions than with OCO-2…" Perhaps you mean OCO-2 inversions have deeper net sinks in comparison to IS? Indeed, it was changed.

Figure A6. Panel titles do not agree with the caption labels. E.g. first panel shows North America, but caption says Boreal North America, and the second panel shows North Trop South America but the caption says temperate North America. Modified

Line 554. Clarify what you mean with "drop off sinks" Modified with: For instance, it seems that the sink decreased for 2018

Line 555. This is an example where you can be more concise and to the point: "…is balanced by the Tropical Asia (North and South) where net fluxes go from sources to sinks."

You can write this as: "…is balanced by sinks in Tropical Asia (North and South) " Changed

Line 555. An example of a sentence with inconsequent use of tense: "2015 was a large net sources of carbon (due to intense fires) while 2016, 2017 and 2018 are deeper sinks with IS over Northern Tropical Asia and sinks with OCO-2 over Southern Tropical Asia. "

I would write it like: In 2015 there were large net carbon sources (due to intense fires), while in the other years there were larger sinks over Northern Tropical Asia (with IS) and Southern Tropical Asia (with OCO-2) Changed

Line 557. An example of unnecessary sentence. “At the same time, posterior fluxes in Europe are anti-correlated with posterior fluxes over Northern Tropical Africa.” The sentence that follows already makes the point of anticorrelations between Europe and Northern Tropical Africa. Deleted

Line 559. “where the post ENSO period has smaller sources in Northern Tropical Africa linked with smaller sinks in Europe (Fig. 11). ”

I don’t think this is true. The deepest sink for Europe appears in 2017, not in 2015. We removed this sentence.

Line 560. “more in line with…” more in line than what? IS inversions? Be clear please. Done

even though the years of study were before 2015.

Line 561. Remove “carbon per year”. It is redundant Done

Line 563. “no estimate that can be refuted at present. ” Do you mean: there is no reliable benchmark for comparison? Yes, we modified it.

Line 565. by Houweling et al. (2015) Done

Line 568. “We can see that our inversions here are within the estimates observed in the study of Peiro et al. (2022).” You can skip this sentence. You already say something similar in lines 566-567. Done

Line 570. Another example of a some text that is very difficult to read

“Our re-balanced priors give the deepest sink in 2017 (in 2016 for CMS prior) which is observed as well in the posteriors net fluxes using OCO-2 and it is in opposition of the OCO-2 inversions of Peiro et al. (2022) which have deeper sinks in 2016. This is due to stronger fire emissions in 2017 compared to the other years balanced with the respiration, and the differences between the two studies could be due to the re-balanced respiration. ”

Can you not write something like:

A major difference between this study and Peiro et al. (2022) is that the rebalanced priors and posterior fluxes provide the largest sink in 2017, as opposed to 2016. This is likely a consequence of the larger fires and the subsequent rebalanced respiration that was derived in this study. Changed

Line 575. “but an agreement across priors within each observational constraint.” What agreement across priors are you referring to? Both types of constraints (OCO-2 and IS) use the same set of prior emissions. So I don’t understand this sentence.

Maybe if you write it like this it becomes clearer what you try to say:

Between all inversions the largest differences in fluxes appear between IS and OCO-2 constraints. However, across the different fire emissions we observe a split; on one hand inversions using FIREmo are similar to FIRE4, while inversions using CMS are more similar to FIRE3. That means fires have a larger impact on the posterior solution than the rebalancing of prior NEE to match the global AGR. Changed

Line 579. “are balanced with higher sources for the other regions that have net sources, regions mainly over the Tropics” please rephrase.

“are balanced with larger sources in other regions, mainly over the Tropics.” Changed

Line 586. “lag between flux in the Tropics and observation by the in situ network” Why is there such a lag in the tropics and not in the extra-tropics? Is that because the distances are longer between the measurements sites and the major source/sinks regions? Please explain in the main text.

We added the sentence: The number of in situ observation is particularly low in the tropics compared to the extra-tropical southern and northern hemispheres (Fig. 2 of Peiro et al., (2022)). One possible explanation is the lag between flux in the Tropics and observation coverage by the in situ network, which could be aliasing flux signals in time, though this hypothesis is difficult to test.

Line 589. "FIREMo and FIRE4 drop off for 2017 but FIRE3 driven fluxes do not."

Please be clear. Do the fire emissions become smaller or the inferred NEE fluxes become smaller? Which region?

Overall I think section b needs to be edited. The main points should be addressed in a more clearer and concise way. The current summary of results is very extensive and long which makes it difficult for the reader to extract the key points.

The key points from section b that needs to be highlighted in a more concise manner:

- Between IS and OCO-2 inversions there are persistent differences in posterior NEE

- Some of these differences are caused by differences in data coverage, lag between flux and observation, cloud fraction, etc.

- Larger sinks with OCO-2 in North America and Europe, while larger sinks with IS in Asia.

- Independent of observational constraints: the sinks in the tropics are generally smaller while there are larger net sinks in the NH Ext.

- Independent of observational constraints: Generally smaller sinks during El Nino in the tropics.

We rewrote the section for better clarification and higlighted the key points at the end of the section.

Line 638. "Additionally, for the 2015-2018 period, the posterior biases were ∼ 7 ppb underestimated TCCON values while the priors were ∼ 13 ppb overestimated TCCON values." Change to: "In comparison to TCCON, for the 2015-2018 period, the posterior biases were underestimated by 7 ppb, while the priors were overestimated by 13 ppb" Done

Line 644. "over the Northern latitudinal" change to: "over the Northern hemisphere" We modified this paragraph with a new figure evaluating the mixing ratio for each TCCON site. See comments Reviewer #2.

Line 654. "In Southern Hemisphere, MOre prior has smaller biases than GFED4re." Are these differences significant at all? All lines seem to be on top of each other in Figure 14. Are the differences significant in comparison to the measurement precision of TCCON CO_2? Please elaborate on this. We have a new figure for the TCCON evaluation which give an evaluation against each TCCON sites. This new evaluation allows a better visualization of the differences between the inversions. As mentioned in the manuscript line 210 "The global monthly means of the total column CO 2 measurements have accuracy and precision better than 0.25% (less than 1ppm) relative to validation with aircraft measurements". Wunch et al., (2010) have shown that any differences with magnitudes less than 0.4 ppm could be attributable to TCCON station site-to-site biases. The evaluation at the Ascension Island site, for instance, biases for ISCMS, IS3re, IS4re and ISMOre are respectively of -0.61, -0.63, -0.56 and -0.55 ppm. The differences and biases cannot be attributed to the TCCON measurement precision. For this tropical site, a bias reduction of 0.8 ppm (0.6 ppm) with ISMOre is obtained compared to IS3re (ISCMS).

Table 6. I suggest to rename the labels to FIRE3, FIRE4 and FIREMo We removed the table as we instead annotated the biases directly with each TCCON site evaluation to reduce paper length and for more clarity.

Line 694. Please discuss the discrepancy between your inversion study and Nechita-Banda et al. regarding the 2015 Indonesian emissions. They found emissions of 0.5 PgC, which is not only more than GFED4, but also more than the GFED3 estimate reported in Fig. 9 of your study.

What is exactly mentioned in the paper of Nechita-Banda et al., 2018 is "Our estimates of CO emissions can be used to quantify the release of gaseous total carbon emissions to the atmos- phere (which includes CO_2, CO, CH_4 and NMVOC). For this conversion, we need to use biomass burning emission factors, which are quite uncertain. Based on our range of results and a range of emission factors available in the literature [14,15,39], we find that a range of 0.35–0.60 Pg C was emitted from the 2015 fires in Indonesia and Papua. ". They did not find emissions in 2015 of 0.5 PgC but a range of 0.35-0.60 PgC. Our CO_2 fire emissions for the southern tropical south Asia give exactly 0.37 PgC for FIRE3, 0.33 PgC for FIRE4 and 0.35 PgC for FIREMo (which can be seen in Fig. 9 of our pre-print). We calculated the emissions for 2015 over the same Indonesia and Papua region of Nechita-Banda, and we found fire emissions of 0.41 PgC for FIRE3, 0.37 PgC for FIRE4 and 0.39 PgC for FIREMo. These fire estimations are included in the range found by Nechita-Banda et al. There is, consequently, an agreement with our CO2 fire estimates and those found in Nechita-Banda et al., (2018). We consequently added: Nechita-Banda et al., (2018) converted their CO fire emissions in CO$_2$ emissions using emission factors and estimated that a range of 0.35-0.60 PgC was emitted in Indonesia and Papua from the 2015 fires. We calculated our fire CO$_2$ emissions over the same region and found 0.41 PgC, 0.37 PgC and 0.39 PgC for FIRE3, FIRE4 and FIREMo respectively. Our fire CO$_2$ estimates are hence in agreement with those found by Nechita-Banda et al., (2018).

Line 695. If GFED4s is able better capture small fires then please explain why GFED3 predicted larger emissions for Indonesia in 2015. Is this related to higher fuel loads in the older model? We found lower fire emissions for southern tropical south Asia with GFED4 compared to GFED3 in 2015. Even though our study period is different, this low GFED4 fire emissions compared to GFED3 were also found in the study of Shi et al., 2015. A possible explanation could indeed come from the CASA biogeochemical model predicting higher biomass densities than with the new version. Additionally, fuel loads in GFED4 for savanna and grassland have been found lower than measured in the field. We added in the manuscript: As mentioned previously, we know that GFED4.1s has information of small fires compared to GFED3 which allow better accuracy particularly over the Tropics where peat fires are important. However, we can see lower FIRE4 emissions than FIRE3 for southern tropical south Asia, similarly to what Shi et al., (2015) have found for the 2002-2012 period. A possible explanation could be that the CASA biogeochemical model of GFED3 predicts higher biomass densities than with the new version used in GFED4. Validation against fuel loads measured in savanna and grassland field have been found higher than with GFED4 (Randerson et al., (2012), Giglio et al., (2013)).

Line 702. “where IS4re and ISMore have sources of carbons compared to the IS constrained with the GFED3 fire, showing then higher net sources with GFED4 and MOPITT than with GFED3 fires”

change to “where IS4re and ISMore derive carbon sources in contrast to IScms that derives a carbon sink with GFED3 fires.” We have changed to: This is particularly shown for the IS inversions where IS4re and ISMore have higher net sources of carbons compared to the IS constrained with GFED3 fires.

Line 704. "the CO2 posterior emissions using MOPITT CO information were able to catch the seasonality"

Please clearly state that you referring here to the CO2 fire emissions and not NEE. The GFED emissions also show seasonality. Is there independent proof MOPITT derived fires show a better seasonality? I would think the smoke could have hampered the MOPITT observations just as much as the MODIS observations. Please elaborate on this. We changed with: Moreover, FIREMo was able to catch the seasonality of fires over southern tropical Asia during the El Nino event, compared to the other priors using GFED inventory. As discussed in NechitaBanda et al., (2018) and van der Laan-Luijkx et al., (2015), GFED4 does not capture fire seasonality due to the use of burned area, compared to GFAS. We already mentioned this line 416 of the pre-print. The burned are may be more sensitive to the initial stages of the fire than the continued burning. GFAS based on the active fires product of MODIS seems to capture fire seasonality compared to GFED (NechitaBanda et al., (2018) and van der Laan-Luijkx et al., (2015)). Figure A4 of the pre-print shows a similarity in seasonality between our FIREMo emissions and GFAS from van der Laan-Luijkx et al., (2015), with a fire peak later than GFED4. However, in both GFED and GFAS method (and similarly for MOPITT), the detection of fires underneath clouds and below the canopy is difficult. But, FIREMO emissions, compared to FIRE3 and FIRE4, has the advantage of combining optimized fire emissions with local observations.

Line 706. "It is thus important to include CO fire emissions over this region to improve estimates and constrain CO2 NEE and Fire emission with both OCO-2 and IS data constraints "

I don't think this sentence covers your methodology correctly. I suggest to rephrase it differently: "It is thus important to use CO observations to constrain estimates of CO2 fire emissions, and subsequently constrain NEE with OCO-2 and IS observations" Changed

Line 709. "finer enough" to "fine enough" Done

Line 710. "Additionally, the emission factors used in the emission ratio are characteristic of vegetation type but are not dependent of spatial or temporal scales. "

I think you try to say that emission factors lack spatial and temporal variability to account for the full dynamics range of combustion characteristics. That is different than saying “not dependent of”. We changed the sentence with this suggestion

Line 713. “between the priors using CO fire emissions and the other prior fire emissions ”

I don’t understand this comparison. All your experiments use fire emissions. Do you mean comparing optimized fire emissions (FIREmo) with non-optimized fire emissions (FIRE4)? Changed with: This could explain the differences observed over some regions of the Tropics between FIREMo and the other prior fire CO2 emissions.

Line 713. “Further works are needed ” to “Further work is required” Done

Line 716. “this tropical region”. This is a new paragraph, so which region are you talking about now? We changed: The data used to constrain inversions is very important. We could see up to 0.4 PgC/yr differences between OCO-2 and IS inversions in tropical regions. This bring us to the importance of the data assimilated in the inversions but also about the priors used in the inversions concerning the different sectors (fire and terrestrial emissions).

Line 730. “The IS results suggest a very strong sink in North Asia ” Do you think this is mostly an inversion artefact due to low data coverage here? As mentioned by Reviewer #2, this paragraph is more a result than a discussion. We consequently removed it. The answer to the question is yes. The disagreement between the OCO-2 and in situ inversions might be driven by the differences in the amount of data assimilated since both inversions have the same transport model and inverse setup. We know that there are fewer in situ than OCO-2 observations above northern Asia, and particularly above the boreal forest of Eurasia, which is an important area for sources and sinks of atmospheric CO2 (Houghton et al., 2007; Siewert et al., 2015).

Line 746. “OCO-2 measurements are globally distributed, but seasonally varying coverage.”, Another difference is that OCO-2 represents a column density as opposed to a concentration in the lower boundary layer. Indeed, and this difference was already mentioned. But for clarity, we changed the sentence: Certainly some of this mismatch is due to sampling differences, as most of the in situ measurements assimilated here are taken in the atmospheric boundary layer in the Northern Extra-Tropics, whereas OCO-2 measurements are globally distributed, but seasonally varying coverage.

To:

Part of this discrepancy is certainly due to: (i) most of the in situ measurements assimilated here are taken in the atmospheric boundary layer while OCO-2 represents a column density; and (ii) most of the in situ measurements are in the northern extra-tropics, whereas OCO-2 measurements are globally distributed, but with seasonally varying coverage.

Line 771. Looking at the posterior fluxes and the TCCON comparisons in Table 6, I hardly see any differences in performance between MOre and the other inversions (GFED3re, GFED4re and CMS). So the added value of optimizing fire emissions before optimizing NEE is not very apparent. On the contrary, your results seem to be very insensitive to the optimized fire emissions. This outcome should be presented much clearer in your discussion and conclusions. We already have modified the sentence:

Regarding the question of the importance of the prior and the question of which prior could do better than the other, we have seen through the results and the evaluation, than no simulation is better than the other on average. Even if the biases seem to have been reduced with FIREMO for certain sites (such as Ascension island for instance), they are in the same order as the other a priori biases for other site. On average and and overall, the added value of optimizing fire emissions before optimizing NEE is not very apparent. Our results seem, overall, to be very insensitive to optimized fire emissions. Philip et al., (2019) performed simulation experiments with different NEE priors, and concluded that posterior NEE estimates are insensitive to prior flux values. But they found large spread among posterior NEE estimates in regions with limited OCO-2 observations. Our results suggesting that OCO-2 inversions are relatively insensitive to prior in most regions, are consistent with Philip et al., (2019), and not only for OCO-2 inversions but also for IS inversions.

We also added in the conclusions: The added value of fire emission for NEE optimization is not apparent. Our results seem hence to be very insensitive to optimized fire emissions.

Line 776. "We found that a priori CO_2 flux uncertainties are substantially reduced when matching the NOAA AGR as well as CO/CO_2 ratio but not strong enough compared to a re-balanced GFED and GFED4.1s NEE, and suggest hence for future work the development of joint CO-CO_2 inversions with multi-observations for stronger constraint in posterior CO_2 fire and biospheric emissions. "

This is a key sentence as it wraps up your paper. However, even after reading the paper I have difficulty to fully understand it.

I tried to rephrase it in three separate sentences. Is my interpretation correct?

"We found that CO_2 fluxes are more robust if the NEE and fire emissions are rebalanced in order to match the NOAA AGR as well as the satellite-based CO constraints. However, a more reliable NEE is obtained if we utilize in situ and satellite-based CO_2 constraints. This opens new avenues for future research for the development of a joint CO-CO_2 inversion framework that uses multiple streams of data to improve the fire and biosphere emissions."

We changed the sentence with: We found that CO_2 fluxes are more robust if the NEE and fire emissions are rebalanced in order to match the NOAA AGR. However, a more reliable NEE is obtained with the assimilated data, using either in situ or satellite-based CO_2 constraints. This opens new avenues for future research for the development of a joint CO-CO_2 inversion framework that uses multiple streams of data to improve the fire and biosphere emissions.

References:

Ramo, R., Roteta, E., Bistinas, I., van Wees, D., Bastarrika, A., Chuvieco, E., and van der Werf, G. R.: African burned area and fire carbon emissions are strongly impacted by small fires undetected by coarse resolution satellite data, Proceedings of the National Academy of Sciences of the United States of America, 118, 1–7, https://doi.org/10.1073/pnas.2011160118, 2021

Shi, Y., Matsunaga, T., Saito, M. et al: Comparison of global inventories of CO2 emissions from biomass burning during 2002-2011 derived from multiple satellite products, Environmental Pollution, 206, 479-487, http://dx.doi.org/10.1016/j.envpol.2015.08.009, 2015

Houghton, R. A., Butman, D., Bunn, A. G., Krankina, O. N., Schlesinger, P., and Stone, T. A.: Mapping Russian forest biomass with data from satellites and forest inventories, Environ. Res. Lett., 2, 4, https://doi.org/10.1088/1748-9326/2/4/045032, 2007.

Siewert, M. B., Hanisch, J., Weiss, N., Kuhry, P., Maximov, T. C., and Hugelius, G.: Comparing carbon storage of Siberian tundr and taiga permafrost ecosystems at very high spa- tial resolution, J. Geophys. Res.-Biogeo., 120, 1973–1994, https://doi.org/10.1002/2015JG002999, 2015.