This manuscript presents inverse model estimates of global and regional CO2 fluxes over the last two decades. The inverse model is based on a single transport model assimilating observations from 50 sites. A series of 16 model simulations is conducted by varying the prior fluxes and prior and observational errors. Results are evaluated against independent aircraft data. The authors found that the ensemble mean of 16 optimized fluxes outperformed individual model outputs. The spread of flux estimates from these 16 model simulations is considered as the uncertainty of the estimated fluxes.

We sincerely thank the reviewer for carefully reading our manuscript and providing us important feedbacks. We have tried our best to address them. Please find answers below each comment. Please find our detailed replies in black to each comment in grey.

General comments

The manuscript presents a detailed study focusing on the inverse model estimation (using a single model) of CO2 fluxes on a global scale for two decades. Therefore, the paper is worthy of publication in ACP after addressing the concerns listed below.

Thank you very much for appreciating our study.

Authors should present the novel aspect of this manuscript. This study uses a single inverse model and conducts a series of model simulations by changing model components, keeping the same observational dataset. Many model intercomparison projects (TransCom and GOSAT and OCO-2 inverse model intercomparisons) address the same aspects by including different transport models but by changing individual model components. Calculating the ensemble mean and spread using a single transport model is not the right way of quantifying the mean and uncertainty in CO2 flux estimates (by not accounting for transport errors).

The novelty of this work is to understand the impact of prior fluxes, and uncertainties (model data uncertainty and prior flux uncertainty) on the estimate of posterior fluxes at the global and regional scale. We agree that the single model transport is not ideal, but please note that this study do not aims to give a full flux estimation uncertainty range, including that due to model transport. We aim to estimate the best CO_2 flux with our transport and provide uncertainty on the best-estimated flux. The transport errors are

already explored extensively in the TransCom, GOSAT, and OCO-2 model intercomparison projects. We evaluate fluxes using independent flux estimation such as RECCAP and compare the simulation of a posteriori flues with independent aircraft data like GCP. These are mentioned in the Abstract.

The MIROC4-ACTM model transport quality has been checked separately using multiple tracer simulations as discussed in the methods section. The performances for inter-hemispheric gradient and vertical transport in the upper troposphere and lower stratosphere are reasonable. Nevertheless, no uncertainty is given to our flux estimation system.

We have now revised the manuscript and do our best to highlight the novelty of the work.

To investigate the impact of different modeling components such as model transport, priors, and specification of uncertainties, there could be other systematic approaches, such as designing a series of simulations and quantitatively assessing the uncertainty components. For example, see Basu et al. (2018) and Philip et al. (2019). More rigorous experiments are required if this manuscript intends to assess the spread from priors and prior/observation uncertainties.

We agree approaches in Basu et al. (2018) and Philip et al. (2019) are good, but for full scale transport error uncertainty estimate. However, the uncertainty we estimate is different from that estimated in Basu et al. (2018) and Philip et al. (2019). Our aim is to estimate uncertainty due to choices of prior fluxes and representation of model data and flux uncertainties. It is impossible to estimate the role of inverse model input parameters without running a single model with a number of different choices of inverse model parameters (e.g., prior flux uncertainty, model data uncertainty etc.); hence we use a well-tested single transport model. As stated earlier we compare our estimated fluxes with regional fluxes from independent studies (RECCAP).

Randomly selecting two different terrestrial biosphere models (TBMs) or ocean models is insufficient. Otherwise, reconsider the focus of the manuscript. This study mainly tests land flux scenarios with and without interannual variability (IAV) (CASA versus VISIT). They should consider using different TBMs as priors (diagnostic/prognostic/with and without IAV etc.) with significant regional differences. That can lead to a reasonable spread in the optimized fluxes. Also, how about conducting a sensitivity test by artificially imposing zero net annual flux in the VISIT model?

While making choices on prior flux selection we did look in to the VISIT land and JMA ocean fluxes in comparison with typical DVGM simulations (GCP's TRENDY) and other ocean observation based flux products used in IPCC AR6, respectively. As you can see from the two plots below our Prior fluxes (CASA and VISIT; Takahashi and JMA) are fairly well encompassing the other available products. Therefore, we did not believe we needed more variety in our prior fluxes.

Figures 4, 5, 6 and 7 in the article are testimony that the inversions are able to bring a posteriori fluxes to a common value from both the prior flux extremes.

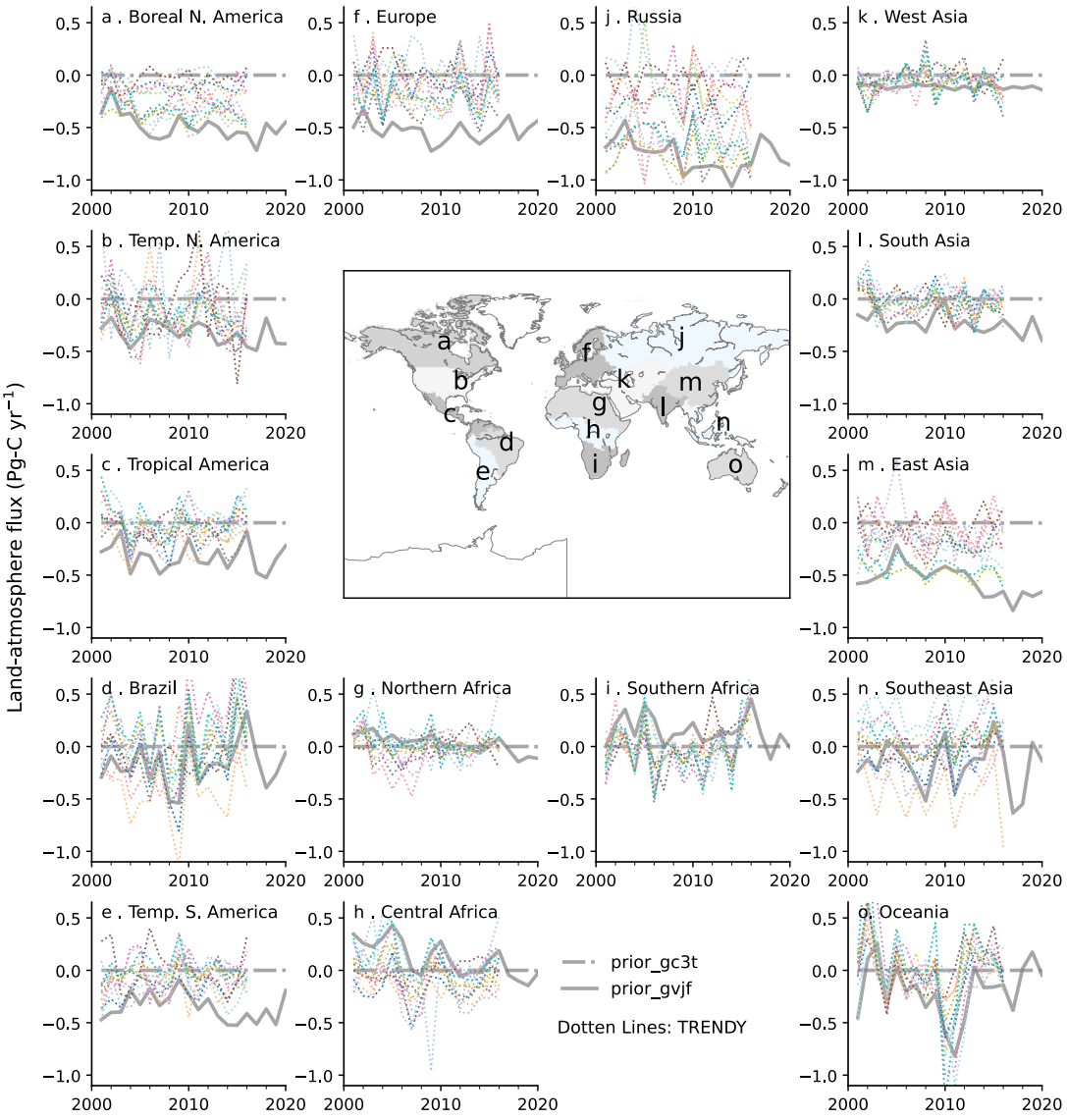

Figure 1. Comparisons of VISIT and CASA prior cases with TRENDY DGVMs (Sitch et al., 2015; https://doi.org/10.5194/bg-12-653-2015) for 15 land regions.

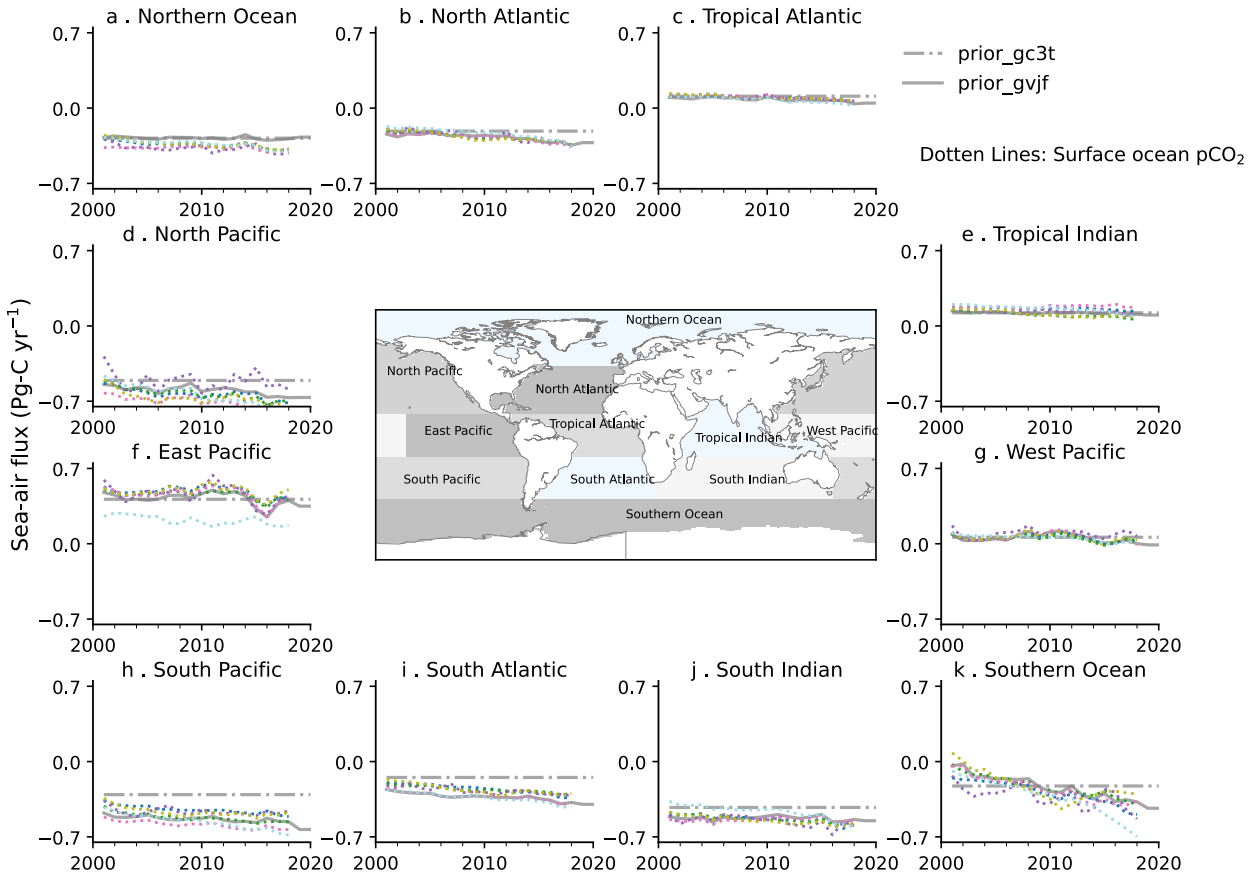

Figure 2. Comparisons of JMA and Takahashi prior cases with pCO2 observation-based products (Fay et al., 2021; https://doi.org/10.5194/essd-13-4693-2021) for 11 ocean regions.

The manuscript should be written more carefully, especially the introduction and conclusion sections. There are many empty/loose sentences, no connection between paragraphs, introduction not providing any motivation of the paper (it also discusses unrelated aspects), grammatical mistakes, etc., throughout the manuscript. See some of the corrections in the technical-correction section below.

Thank you for pointing this out here and suggesting numerous corrections below. The reviewer#2 is also very kind in reading the manuscript line by line and pointing us to all corrections that are needed. We have made our best effort to incorporate all suggestions and revise the manuscript.

Specific comments

Line 17-21: These two sentences are not connected. You state that model errors and insufficient observations lead to uncertainties in regional flux estimates. However, it is unclear how you address these with your simulations using a single model. State clearly what uncertainty component you are addressing here in this article.

We have revised the two sentences as

"However, the uncertainties in the regional flux distributions remain unconstrained due to the lack of high-quality measurements, uncertainties in model simulations, and representation of data and flux errors in the inversion systems. Here, we assess the representation of data and flux errors by using a suite of 16 inversion cases derived from a single transport model (MIROC4-ACTM) but different sets of a priori (bottom-up) terrestrial biosphere and oceanic fluxes, as well as prior flux and observational data uncertainties (50 sites) to estimate CO_2 fluxes for 84 regions over the period 2000-2020."

Line 26-28: This is just a general statement. Need more clarity here: "Interannual variability and seasonal cycle in CO2 fluxes are more consistently derived for different prior fluxes when a greater degree of freedom is given to the inversion System".

We have slightly revised the sentence as

"Interannual variability and seasonal cycle in CO_2 fluxes are more consistently derived for two distinct prior fluxes when a greater degree of freedom (increased prior flux uncertainty) is given to the inversion system."

Line 28-29: In line 261, you mention that fluxes are evaluated with aircraft observations. Are you using surface data as well? "…evaluated the inversion fluxes using independent aircraft and surface measurements not used in the inversion".

Yes, surface data are also used, say in Fig. 12 (bottom row)

Line 28-29: Good if you can make it more quantitative, i.e., add some summary statistics or so: "which raises our confidence in the ensemble mean flux rather than an individual inversion"

We consider this remark and revised the sentence as

“We have further evaluated the inversion fluxes using meridional CO_2 distributions from independent (not used in the inversions) aircraft and surface measurements, suggesting that the ensemble mean flux (model-observation mean±1σ standard deviation = 0.3±3 ppm) best suited for global and regional CO_2 flux budgets than an individual inversion (model-observation 1σ standard deviation = 0.35±3.3 ppm).”

Line 31: It seems like an empty/loose sentence: “Differences between 5-year mean fluxes show promises and capability to track flux changes under ongoing and future CO2 emission mitigation policies.”

We have slightly revised the sentence for clarity as
“Using the ensemble mean fluxes and uncertainties for 15 land and 11 ocean regions at 5-year intervals, we show promises and capability to track flux changes toward supporting the ongoing and future CO_2 emission mitigation policies.”

Line 36-38: Cite IPCC report.

Canadell et al. is the referring to the Chapter of IPCC AR6 which has assessed the TCRE etc. Cited again in the next sentence, instead of this general sentence.

Line 44: Be very clear (solutions to …?): “The sinks on the land and ocean constitute a major component of nature-based solutions”.

We modified the sentence as
“The sinks on the land and ocean constitute a major component of nature-based solutions to mitigate the rise in CO_2 concentration, as discussed in the IPCC AR6 (Canadell et al., 2021)”

Line 45-46: Cite proper references to support the statement.

The paragraph is rearranged and shifted to Line#57. We have added a reference to Kondo et al. (2020) for this sentence.

Line 45-53: In this paragraph, mention global flux uncertainty first, and then note the regional issues, with some additional details. That is, lines 45-46 should come after line 53.

Rearrangements of the text is done as per your suggestions.

Line 55-69: It is not clear why you need this paragraph. "However, the impacts of biases in FFC emissions on inversion estimated CO2 fluxes remained relatively unexplored". Are you exploring this aspect in this paper? Moreover, this paragraph is written poorly.

The aim of this introduction on FFC aliasing effect to give the background of the discussions related to the flux trends for East Asia; Fig. 6 and Fig. 10. Some revisions are now made based on reviewer#2's suggestions and we have added some text in the previous paragraph to link the topics. Hope the paragraph reads better now.

Line 70-73: I don't quite understand this statement! Who provides the metric, what is that metric? What's the meaning of "metric for evaluation of regional fluxes should be evaluated"? Clarify.

We have now revised this paragraph. The 2nd sentence is deleted in the process. Hope the paragraph read well now. The revised paragraph is

"The GCP annual updates of inversions provide a metric for evaluating inversions using independent measurements, mainly from the aircraft campaigns (e.g., Friedlingstein et al., 2020). Evaluation of predicted fluxes from model-data differences may not be straightforward due to the underlying assumptions of a flux inversion system, e.g., for flux correlation lengths or the radius of influence for the measurements, observational data uncertainty, prior flux uncertainty (Baker et al., 2010; Chevallier et al., 2007; van der Laan-Luijkx et al., 2017; Miyazaki et al., 2011; Niwa et al., 2017; Rodenbeck et al., 2003), while the data assimilation system will fit the model concentrations to the observed values. For example, a model-observation difference within ±1 ppm and/or vertical concentration gradient simulation within 1-σ standard deviation of the observed gradient resulted in more than 1 PgC yr^{-1} flux differences between models at regional or sub-hemispheric scales (Gaubert et al., 2019; Stephens et al., 2007; Thompson et al., 2016). Another way of improving our knowledge about uncertainties in regional flux estimations is to employ multiple types of datasets from both bottom-up and top-down modelling systems (Ciais et al., 2021; Kondo et al., 2020), which we have adapted here for checking the regional inversion fluxes, in addition to the GCP like evaluation using independent aircraft data."

Line 71-73: Is this something new? "…should be evaluated using a new transport model simulation of the predicted fluxes, not using the assimilated CO2 field". Be clearer with sufficient details. Most evaluations in current published works are based on model simulation of optimized fluxes. For evaluation, using a

different transport model than the one used in the inversion (as a forward model) is advantageous (not sure if this is what you mean here). Also, are you exploring this in this manuscript/study?

No, unfortunately this could not be explored in details and remained as a hypothesis. Thus, we have revised this paragraph as mentioned above by cleaning up texts.

Line 73-81: I'm lost here. From re-reading this, I understand that the assessment of the spread of optimized fluxes obtained by conducting multiple simulations using different model inputs is a better way of quantifying the uncertainty than simply evaluating the optimized CO2 concentrations against independent measurement data. Revise the entire paragraph to be more apparent.

Thank you. We have revised the paragraph as a whole, and added at the end as discussed in previous comment (Line 70 – 73)
"Another way of improving our knowledge about uncertainties in regional flux estimations is to employ multiple types of datasets from both bottom-up and top-down modelling systems (Ciais et al., 2021; Kondo et al., 2020), which we have adapted here for checking the regional inversion fluxes, in addition to the GCP like evaluation using independent aircraft data."

Line 82-85: These uncertainty sources have been investigated previously. Cite some of those critical studies here.

We have cited : (Basu et al., 2018; Patra et al., 2005a; Philip et al., 2019; Qu et al., 2021; Wang et al., 2018).

Line 92-95: This statement is not correct: "Such intercomparisons used single inversion from different modeling groups and provided the range in CO2 flux uncertainty due to differences in transport models.". These intercomparisons assessed uncertainty arising from different model components, not just the transport model differences. For example, see Crowell et al., 2019 and Peiro et al., 2022.

We are aware of these publications (Crowell et al. 2019 is already cited; Peiro et al. used fluxes for an extended period) which makes assessments of regional fluxes estimated by inversions using OCO-2 data but none of both give separate assessments of prior flux uncertainty vs data uncertainty on the inversion results, for example. In fact, it is impossible to estimate the role of inverse model input parameters without running a single mode with a number of different choices of inverse model parameters (e.g., prior flux uncertainty, measurement data uncertainty etc.).

Following your and Reviewer#2 suggestion we have modified the statement as

"Such intercomparisons used single inversions from different modelling groups and provided the range in total CO_2 flux uncertainty due to the choices of prior fluxes distribution, prior flux uncertainty, observational data uncertainty, and the model transport uncertainties."

Line 123-124: This sentence is not clear to me.

The sentence is revised as

"Reasonably good model transport in MIROC4-ACTM enables us to use any mismatch between observation and simulations to estimate the land and oceanic fluxes using the inverse modelling technique (details in Section 2.4)."

Line 128 and 129: Just "used" not "simulated" (?): "… is simulated using …"

Corrected as per your suggestion.

Line 135: "…downscaled to 3-hourly time intervals…": Mention how you downscaled; which variable used; and cite proper literature.

Thank you for pointing this out. The sentence is revised as "The CASA and VISIT monthly-mean fluxes are downscaled to 3-hourly time intervals by redistributing respiration and gross primary production (Olsen and Randerson, 2004) using JRA-55 meteorology, i.e., 2m air temperature and incoming solar radiation at the earth surface"

Line 136: Double-check if it is version 4.1? "…fire emissions are used from GFEDv4s (van der Werf et al., 2017…".

Thank you for catching the mistake. Yes, we have used v4.1s

Line 145: Complex notations: gc3t and gvjf. What is "3" and "t" in gc3t?

Agreed, but we created a 4 lettered name to accommodate all 4 fluxes, as given in Eq. 1. We have slightly rearranged for better clarity about how the notations are formed: g: GridFED FFC, c3:CASA-3hr, v:VISIT, t:TT09 ocean, j: JMA ocean, and f: Fire

Line 149-150: Revise: "to evaluate the strength of MIROC4-ACTM simulations to derive fluxes consistently". How do you evaluate the strength of simulation? Why did you mention "consistently" here? Fluxes will be derived using the inverse model, so how can you "evaluate the strength of forward simulation"?

Sorry for the unclear information. Our intent was to evaluation inversion strength. So, we have changed "simulations" by "inversions" for clarification. We stated both - consistently (or the lack of it)! We think transport model is key for the fluxes we derive, although CO_2 is an inert species and transport is linear. So, we thought of mentioning the model's name.

Line 159: Cite proper references: "WDCGG websites as appropriate"

Given as : GML/NOAA (https://gml.noaa.gov/aftp/data/trace_gases/co2/flask/) and WDCGG (https://gaw.kishou.go.jp/) websites

Line 162: Is this the grid cell with the observation location? "…nearest grid of observation location at hourly intervals…".

Revised as "the grid point nearest to the observation location" for clarity

Line 164: "These temporal data gaps (1-6 months) are filled using the curve fitting method based on the digital filtering technique". Have you conducted simulations without using curve-fitted data? Why was this data filling necessary?

The matrix inversion system requires data every month of the inversion period. We have checked the fitting program works well when data gap is less than 6 months, as the seasonal cycle is derived by using 6 harmonics.

Line 200-210: How about conducting a simulation with "gpp_v4" along with "ocean PFU = 0.5"? Explain the rationale for selecting different prior error scenarios you considered in this study.

We have stated the rationale as: (sorry without conducting the suggested simulation)
Selection of wide range of PFUs, in the range of 0.5 – 1.0 PgC yr^{-1} the ocean regions and 0.2 – 4.0 PgC yr^{-1} for the land regions allows us to understand about the stability of the inversion system as assess the range of a posteriori fluxes for aggregated sub-continental/basin regions or the land and ocean totals.

Line 234: "High values (FUR towards 100)": If FUR is in percentage, then revise the equation in line 233.

Revised as per your suggestion.

Line 244: Not clear: "… indicative of the observational constraint regional fluxes…"
Line 245: "…we recommend that the spread of ensemble inversions provide more representative estimation of the regional CO2 sources and sinks.". "Spread" represents "a measure of uncertainty", not a "representative estimation of…". Why do you add "recommend" here?

This sentence is revised as "As discussed later in this article, the FUR is only indicative of the observational constraint on the regional fluxes, the spread of ensemble inversions provides a measure of uncertainty of the regional CO_2 sources and sinks."

Line 309 and 311: Revise this sentence: "Hence, the magnitude of biases and RMSE indicates predominantly the accuracy of the predicted fluxes.". Model transport is one of the sources leading to uncertainties in the predicted fluxes.

As a remark. we have revised this sentence as "Hence, the magnitude of biases and RMSE indicates predominantly the accuracy of the predicted fluxes (the errors due to model transport and measurement network are not explored in this study)."

Line 649: "CO2 simulations are derived from three sets of prescribed fluxes: "gc3t", "gvjf", and "ensm".": I'm assuming that the evaluation is conducted for all 16 inversions (?).

All 16 sets are not simulated by model, but we used 3 representative cases, and part of this sentence is now revised for clarity as
"The CO_2 simulations are derived from three sets of prescribed fluxes: "gc3t" (case: ctl_ux4_gc3t in Table 2), "gvjf" (case: ctl_ux4_gvjf), and "ensm" (average of all 16 inversions)."

Lines 709-720: I'm not sure if these details (+ Figure S10) are required in this paper.

Thank you for this suggestion. Review#2 also expressed concerns on this part of the discussion.
We have removed this paragraph and the one before it, and Fig. S10.

Line 775-782: Empty/loose sentences.

We have deleted this final paragraph of the Conclusions, as per your suggestion and also because lines 692-720 in the submitted version are deleted following suggestions from you and Reviewer#2.

Technical corrections

Line 14: Better add "atmospheric" here: "chemistry-transport model (ACTM)".

Line 16: Better avoid text in parenthesis: "regional flux (+ve: source to the atmosphere; -ve: sink on land/ocean)".

Line 21: Move the number of the sites (50) from here to the appropriate part of the sentences: "data uncertainties (50 sites)".

Line 24: Is this "22-33% and 16-18%" for land vs ocean? Not sure this is clear enough here.

Line 25: Not clear what this approximate means here: "best estimations for (approx. 2000-2009)".

Line 52: Revise and add more clarity: "partitioning exists greatly in the … release".

Line 55-56: Revise this sentence: "…because inversion calculations do not optimize…".

Line 90-91: You can write these in a better way: "inversions from … for inversions using … or for inversions".

Line 99: Revise: "observed and model data processing".

Line 100: Avoid capital letter: "the Results and discussion".

Line 155-156: Avoid repetition of "from".

Line 1120: Correct this: "lower panel (b)".

Line 242: Correct: "…West Asia, Northern Africa. The Tropical Indian Ocean…".

Line 252: Correct: "as per analysis".

Line 302-307: Use simple notations. For example, avoid "aircraft" from "x".

Line 308: Correct: "CO2 mixing ratios".

Line 317: Use the term "grid cells".

Line 336: Avoid ".": "3.2. Global totals."

Line 346: Use "mean": "Ensemble means land".

We appreciate your help very much. All of the above corrections are made in the revised manuscript.

.

Line 563: Revise: “It is not easy for us to explain”.

Thank you for the suggestion, we have further scrutinized the Yasunaka et al. paper and added “It is not easy to put forward a hypothesis for the weaker sink in summer than in winter of Northern Ocean, while we can speculate that the atmospheric CO_2 decrease in polar air exceeds compared to the decrease that occur over the surface sea-water and reduced solubility of CO_2 in warmer water. Indeed, Yasunaka et al. (2018) have shown that the Greenland-Norwegian seas and Barents Sea are indeed acts as milder sink of CO_2 (flux = -4 to -5 mmol m^{-2} day^{-1}) during June-August compared to the October-March (flux = -10 to -15 mmol m^{-2} day^{-1}), and the Chukchi Sea and Arctic Ocean show strongest uptake in October. Thus, as whole the Northern Ocean of our study could act as the weakest sink in summer months.”

Line 763: Avoid “Please”.

Line 766: Correct: “is unanimously located”.

Figure 4: Choose a different font that is clearer.

Figure S2: Correct to CO2: “monthly-mean CO fluxes”

All of the above corrections are performed. We appreciate your help very much.

References

Basu, S., et al.: The impact of transport model differences on CO2 surface flux estimates from OCO-2 retrievals of column average CO2, Atmos. Chem. Phys., 18, 7189–7215, https://doi.org/10.5194/acp-18-7189-2018, 2018.

Crowell, S., et al.: The 2015–2016 carbon cycle as seen from OCO-2 and the global in situ network, Atmos. Chem. Phys., 19, 9797–9831, https://doi.org/10.5194/acp-19-9797-2019, 2019.

Peiro, H., et al.: Four years of global carbon cycle observed from the Orbiting Carbon Observatory 2 (OCO-2) version 9 and in situ data and comparison to OCO-2 version 7, Atmos. Chem. Phys., 22, 1097–1130, https://doi.org/10.5194/acp-22-1097-2022, 2022.

Philip, S., et al.: Prior biosphere model impact on global terrestrial CO2 fluxes estimated from OCO-2 retrievals, Atmos. Chem. Phys., 19, 13267–13287, https://doi.org/10.5194/acp-19-13267-2019, 2019.

All of these references are cited in the revised manuscript.