The authors revised their text extensively, but three points have not been properly addressed, the first two being major. The line numbers below are those used in my initial review.

[Response] We appreciate your constructive comments on this manuscript. We revised the manuscript to fully address your comments and suggestions. Detailed point-by-point responses to your comments and related revisions are presented below. The original comments are in black, and our responses are in blue color.

L27 - I understand that the authors do not "conclude that it [one element of the MIP protocol about fossil fuel emissions] is the main source of the underestimation in the true flux errors". But this element in the protocol is a likely explanation that, if correct, would make "the main finding of the study" (authors' words) pretty useless.

[Response] During the previous round of review, we were unsure about the meaning of this comment, "this main conclusion is not a finding but an input to the study". Upon further reflection, we believe that the reviewer is pointing out that the errors we estimated are on the net surface-atmosphere CO_2 fluxes, which is the combined terrestrial biosphere CO_2 fluxes and fossil CO_2 emissions. This is an insightful criticism of the way our results were presented. We have re-oriented the manuscript to emphasize the error estimates are on the net CO_2 fluxes, not just the terrestrial biosphere CO_2 fluxes. We want to emphasize that this reinterpretation does not impact our results, as the ensemble spread of posterior terrestrial biosphere fluxes is identical to that of posterior net fluxes. This is because all OCO-2 MIP models prescribed the same fossil fuel emission estimates and treated them perfectly known values. We revised the manuscript throughout, changing the previously stated "errors in the terrestrial biosphere CO_2 flux estimates" to "errors in the net CO_2 flux estimates" and sentences related to these revisions. The sentences with major revisions related to this are as follows:

L69-71, "This study aims to develop a framework to quantify the errors in **regional net surface-atmosphere CO_2 fluxes (terrestrial biosphere fluxes + fossil fuel emissions)** estimated from an ensemble of inverse models by using airborne CO_2 measurements, transport modeling, and adjoint sensitivity analysis."

L313-318, "Therefore, we can obtain the true errors in the ensemble annual total **net land fluxes** in those areas, err_{f_t} $(= \sigma_{f_t})$, by multiplying the ratio between three-year mean values of $h(err_{f_t})$ and $h(err_{f_e})$ by the ensemble spread of the **annual total net land flux estimates** (err_{f_e}) within the effective areas. The equation can be written as:

$$err_{f_t} = \frac{h(err_{f_t})}{h(err_{f_e})} \times err_{f_e} \quad (12)$$

One thing readers should keep in mind is that the err_{f_e} is identical to the ensemble spread of posterior terrestrial biosphere fluxes because all OCO-2 MIP models used uniform fossil fuel emission estimates and assumed them to be perfectly known."

L467-469, **"**The black error bars denote ± one standard deviation of the **posterior net land fluxes, identical to those of the posterior terrestrial biosphere fluxes.** The error bars in red indicate the newly-estimated range of errors **in the posterior net land fluxes** from this study."

In addition, we removed the first paragraph of the original introduction, which emphasized terrestrial biosphere flux estimates. The **title** of this study is also revised as follows:
"Quantification of regional **net** CO_2 flux errors in the v10 OCO-2 MIP ensemble using airborne measurements"

Moreover, we clearly described the reason for the underestimation of the true flux errors in the regions with high fossil fuel emissions and the implications of our study in the main text. The revised sentences are as follows:

L25-26, "This suggests the presence of systematic biases in the inversion estimates associated with errors in the prescribed fossil fuel emissions common to all models."

L487-501, "Our analysis reveals that the true errors in the ensemble mean of posterior net CO_2 flux estimates is significantly greater than the ensemble spread of flux estimates in five out of seven regions with higher fossil fuel emissions compared to terrestrial biosphere fluxes. Possible explanation for this result is the presence of errors in the prescribed fossil fuel emissions common to all OCO-2 MIP models. OCO-2 MIP models treated fossil fuel emissions as perfectly known values and adjusted terrestrial biosphere and ocean CO_2 fluxes to minimize the difference between the simulated and observed CO_2 concentrations. Thus, if there are errors in the prescribed fossil fuel emission estimates, these errors propagate into the posterior natural flux estimates. The assumption used in the OCO-2 MIP models is, in fact, the one often applied in conventional global atmospheric inverse models as it is considered that the errors in fossil fuel emission estimates are relatively lower than those in natural flux estimates at national scales (4-20%; Andres et al., 2014). However, the emission errors become substantial when considering spatial distribution at model grid scale and temporal variability within a year (Zhang et al., 2016; Gurney et al., 2021). Oda et al. (2023) showed significant impacts of differences in fossil fuel emission estimates on posterior terrestrial biosphere flux estimates near the source regions. OCO-2 MIP models used identical fossil fuel emission estimates and thus their posterior net flux estimates share common biases induced by the errors in the fossil fuel emission estimates. Because these systematic biases are not captured by the ensemble spread of flux estimates, true flux errors exceed the errors computed from the ensemble spread in the main source regions. In addition to this, …"

L604-608, "… This result provides observation-based evidence supporting previous studies (Oda et al., 2023; Wang et al., 2020) that emphasized the impact of fossil fuel emission errors on global atmospheric CO_2 inversions. This finding offers important insights into understanding the sources of errors in current inverse modeling and highlights the need for improving fossil fuel emission estimates and developing inversion methods that account for uncertainties in both fossil fuel emissions and natural fluxes. …"

- Andres, R. J., Boden, T. A., and Higdon, D.: A new evaluation of the uncertainty associated with CDIAC estimates of fossil fuel carbon dioxide emission, Tellus B: Chem. Phys. Meteorol., 66, 23616, https://doi.org/10.3402/tellusb.v66.23616, 2014.
- Gurney, K. R., Liang, J., Roest, G., Song, Y., Mueller, K., and Lauvaux, T.: Under-reporting of greenhouse gas emissions in U.S. cities, *Nat. Commun.*, 12, 553, https://doi.org/10.1038/s41467-020-20871-0, 2021.
- Oda, T., Feng, L., Palmer, P. I., Baker, D. F., and Ott, L. E.: Assumptions about prior fossil fuel inventories impact our ability to estimate posterior net CO_2 fluxes that are needed for verifying national inventories. *Environ. Res. Lett.*, *18*(12), 124030, https://doi.org/10.1088/1748-9326/ad059b, 2023.
- Wang, J. S., Oda, T., Kawa, S. R., Strode, S. A., Baker, D. F., Ott, L. E., and Pawson, S.: The impacts of fossil fuel emission uncertainties and accounting for 3-D chemical CO_2 production on inverse natural carbon flux estimates from satellite and in situ data, *Environ. Res. Lett.*, *15*(8), 085002, https://doi.org/10.1088/1748-9326/ab9795, 2020.
- Zhang, X., Gurney, K. R., Rayner, P., Baker, D., and Liu, Y.-P.: Sensitivity of simulated CO_2 concentration to sub-annual variations in fossil fuel CO_2 emissions, *Atmos. Chem. Phys.*, 16, 1907–1918, https://doi.org/10.5194/acp-16-1907-2016, 2016.

L220 – I respectfully disagree because sub-monthly flux patterns may also affect monthly concentration patterns, depending on sub-monthly transport patterns. This point needs to be documented.
[Response]We agree the reviewer's comment. We included following sentences to carefully discuss the possible impact of variability of sub-monthly flux variations within OCO-2 MIP models on our results in the main text (L587-599):

"This study uses monthly mean posterior flux estimates for the calculation of monthly $h(err_{f_e})$ but posterior flux estimates from each OCO-2 MIP models have different sub-monthly patterns. This could modify the sub-monthly variations in posterior atmospheric CO_2 and affect the ensemble spread of posterior CO_2 concentrations. However, due to absence of information on sub-monthly variations in posterior flux estimates, this study assumed that the contributions of the inter-model variability of sub-monthly flux variations to our monthly mean error quantities ($h(err_{f_e})$ and ERR_{MIP}) are not significant. This assumption is supported by comparing $h(err_{f_e})$ with ERR_{MIP}. ERR_{MIP} resulted from variabilities in not only hourly posterior flux estimates but also transport models. Despite $h(err_{f_e})$ not accounting for the impacts of inter-model variability of sub-monthly flux patterns, the regional mean of monthly $h(err_{f_e})$ (0.44-0.93 ppm), on average, accounts for 58-86% of regional mean of monthly ERR_{MIP} (0.51-1.34 ppm) throughout the analysis period (Figure 5h and Fig. S6). Furthermore, we found that our main results remain robust across the potential range of $h(err_{f_e})$ when it includes the impact of sub-monthly flux variations. For example, if $h(err_{f_e})$ increases, on average, by 0.2 ppm, the ratio of regional mean of $h(err_{f_e})$ to $h(err_{f_t})$ increases from 0.74 [0.61, 0.88] to 0.83 [0.71, 0.96] in midlatitude North America and from 0.59 [0.48, 0.70] to 0.67 [0.57, 0.78] in East Asia throughout the analysis period."

L300 - CMS-Flux is not "most models". The last part of the sentence ("suggesting…") should be deleted.
[Response] We agree that the previous reference does not represent "most models". Instead, we cited Gaubert et al. (2023), which revealed that most OCO-2 MIP inverse models overestimate the observed atmospheric CO_2 concentrations along the African coast during the ATom project, due to potential biases in OCO-2 XCO_2 measurements over northern tropical Africa. We also revised the corresponding sentence accordingly (L362-363) as follows:

"These findings agree with Gaubert et al. (2023), which showing most of the inverse models in v10 OCO-2 MIP have significant errors because of potential positive biases in OCO-2 XCO_2 measurements for this region."

- Gaubert, B., Stephens, B. B., Baker, D. F., Basu, S., Bertolacci, M., Bowman, K. W., Buchholz, R., Chatterjee, A., Chevallier, F., Commane, R., Cressie, N., Deng, F., Jacobs, N., Johnson, M. S., Maksyutov, S. S., McKain, K., Liu, J., Liu, Z., Morgan, E., O'Dell, C., Philip, S., Ray, E., Schimel, D., Schuh, A., Taylor, T. E., Weir, B., van Wees, D., Wofsy, S. C., Zammit-Mangion, A., and Zeng, N.: Neutral Tropical African CO2 Exchange Estimated From Aircraft and Satellite Observations, Global Biogeochem. Cycles, 37, e2023GB007804, https://doi.org/10.1029/2023GB007804, 2023.

In addition, the title now needs "the" before “v10 OCO-2 MIP ensemble”.

[Response] Thank you. We revised the title as follows:

Quantification of regional net CO_2 flux errors in **the** v10 OCO-2 MIP ensemble using airborne measurements