[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.

The authors have worked on my remarks seriously again, but the answer to point "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" is weak. I regret to have to recommend further analysis of this potential weakness of the study.
On a minor side "due to absence of information on sub-monthly variations in posterior flux estimates" is not correct (this information is available somewhere, even though it may not be public).
[Response] We agree with the reviewer's comment. We conducted an additional analysis using various datasets of hourly (or 3-hourly) terrestrial biosphere fluxes to assess how different sub-monthly patterns of posterior flux estimates from OCO-2 MIP models might affect our main results. The revisions based on this additional analysis are as follows (L587-608 in the revised manuscript):

"This study uses monthly posterior flux estimates to calculate monthly $h(err_{f_e})$. However, posterior flux estimates from each OCO-2 MIP model have different sub-monthly patterns, which could modify sub-monthly variations in posterior atmospheric CO_2 concentrations and, in turn, affect their ensemble spread. To examine their potential impact on the results, we conduct an analysis with different publicly available hourly (or 3-hourly) terrestrial biosphere fluxes (Chevallier et al., 2019; Jacobson et al., 2020; Ott et al., 2020; Haynes et al., 2021; Liu and Bowman, 2024) which are from seven OCO-2 MIP prior flux models (Ames, Baker, CAMS, CMS-Flux, CT, OU, and WOMBAT; Table S1). By incorporating the monthly-balanced hourly flux estimates into the monthly posterior fluxes, we generate hourly posterior terrestrial biosphere flux estimates for these seven models. Since the assimilation window for each OCO-2 MIP model ranges from one week to one month, the weekly variations in posterior fluxes may differ from those in the prior fluxes. Nonetheless, with only the monthly posterior flux estimates being publicly available, this approach offers valuable insights into how different sub-monthly patterns of posterior fluxes could affect our main results. Our analysis shows that the regional averages of $h(err_{f_e})$ derived from the monthly posterior flux estimates from the seven models are, on average, within ±10% of the values originally obtained using flux estimates from 10 models for the period 2015–2017, except for Europe (13% lower) (Fig. S12a). When accounting for different sub-monthly patterns of posterior fluxes across models, the regional averages of $h(err_{f_e})$ increase by 10-22% (0.06-0.14 ppm) across six regions, with a 45% (0.23 ppm) increase in Europe. These results suggest that our earlier calculation, assuming identical sub-monthly flux variations, underestimates $h(err_{f_e})$. We further investigate whether our main finding remains robust even if we adjust the original values of $h(err_{f_e})$ using the potential underestimation rate. After making the correction, we found that the ratios of the regional average $h(err_{f_e})$ to $h(err_{f_t})$ increase the most in Europe by 0.14 and only up to 0.07 in the other six regions, as $h(err_{f_t})$ also rises with $h(err_{f_e})$ according to Eq. (8) (Fig. S12b). Moreover, the $h(err_{f_t})$ still exhibits significant underestimation ($p<0.05$) in mid-latitude North America, Europe, East Asia, Southeast Asia, and Australia. This indicates that our main results are robust to the inclusion or exclusion of sub-monthly flux patterns in the calculation of $h(err_{f_e})$."

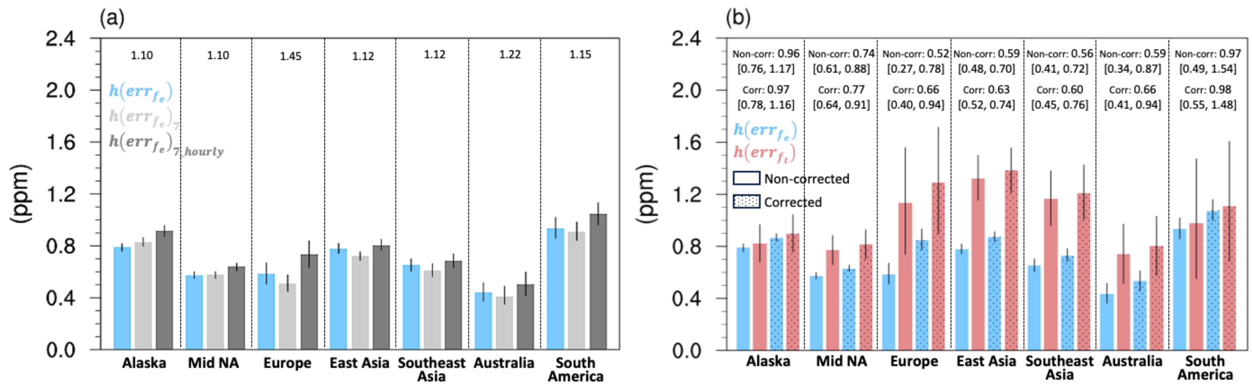

Figure S12. (a) Mean values of monthly $h(err_{fe})$ for each region for the period 2015–2017, derived using different monthly posterior flux estimates from either 10 (blue) or seven ($h(err_{fe})_7$; light gray) OCO-2 MIP models with identical hourly NBE variation information, or from different monthly posterior flux estimates from seven models with different hourly NBE variation information ($h(err_{fe})_{7\_hourly}$; dark gray). The numbers at the top of the panel (a) indicate the ratio of $h(err_{fe})_{7\_hourly}$ to $h(err_{fe})_7$. **(b)** Mean monthly values of $h(err_{fe})$ (blue) and $h(err_{ft})$ (red) for each region over three years, with the dotted bars representing the corrected $h(err_{fe})$, obtained by multiplying the ratio of $h(err_{fe})_{7\_hourly}$ to $h(err_{fe})_7$, and the recalculated $h(err_{ft})$ using these corrected $h(err_{fe})$ values. The numbers at the top of the panel (b) denote the ratio of $h(err_{fe})$ to $h(err_{ft})$. The error bars represent the 95% confidence intervals derived from 1000 bootstrap samples of datasets.

- Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., and Cozic, A.: Objective evaluation of surface- and satellite-driven carbon dioxide atmospheric inversions. *Atmospheric Chemistry and Physics*, *19*(22), 14233-14251, https://doi.org/10.5194/acp-19-14233-2019, 2019.
- Haynes, K.D., I.T. Baker, and A.S. Denning.: SiB4 Modeled Global 0.5-Degree Hourly Carbon Fluxes and Productivity, 2000-2018, ORNL DAAC, Oak Ridge, Tennessee, USA, https://doi.org/10.3334/ORNLDAAC/1847, 2021.
- Jacobson, A. R., Schuldt, K. N., Miller, J. B., Oda, T., Tans, P., Arlyn Andrews, Mund, J., Ott, L., Collatz, G. J., Aalto, T., Afshar, S., Aikin, K., Aoki, S., Apadula, F., Baier, B., Bergamaschi, P., Beyersdorf, A., Biraud, S. C., Bollenbacher, A., … and Zimnoch., M.: *CarbonTracker CT2019B*, NOAA Global Monitoring Laboratory, https://doi.org/10.25925/20201008, 2020.
- Liu, J. and Bowman, K.: Carbon Monitoring System Carbon Flux Land Prior L4 V3, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Last access: 17 June 2024, 10.5067/1XO0PZEZOR1H, 2024.
- Ott, L.: GEOS-Carb CASA-GFED 3-hourly Ecosystem Exchange Fluxes 0.5 degree×0.625 degree V3, Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, MD, USA, https://doi.org/10.5067/VQPRALE26L20, 2020.

Please also review the English of the sentence "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 XCO2 measurements for this region.".

[Response] We revised the sentence (L364-365 in the revised manuscript) as follows:

"These findings agree with Gaubert et al. (2023), which shows that most inverse models assimilating OCO-2 XCO_2 retrievals tend to overestimate the net carbon sources in this region because of potential positive biases in the OCO-2 retrievals.