Referee #2

We would like to thank the anonymous referee for his/her comprehensive review and valuable suggestions. We have made revisions based on the referee's suggestions and have responded to all comments point by point. The page and line numbers of all revisions are referenced to the revised manuscript. References related to the responses are listed in the end of this document.

Summary:

In this work, the authors conduct atmospheric CO_2 inversions to estimate global NEEs using OCO-2 and OCO-3 XCO_2 retrievals and the Global Carbon Assimilation System, version 2(GCASv2). Three sets of experiments have been designed by the authors to evaluate the impact of using different OCO XCO_2 observations to constrain the posterior carbon fluxes: using OCO-3 XCO_2 only; using OCO-2 XCO_2 only; and using OCO-2 & OCO-3 XCO_2 combined. The overall results suggest using combined OCO-2&OCO-3 XCO_2 retrievals can yield better consistency when compared with in-situ observations while using OCO-3 XCO_2 retrievals alone presents largest biases. The results and discussion reveal some interesting patterns in global and regional NEEs across different experimental setups and provided some insights on the choice of satellite observations to constrain global NEEs, but lack in-depth discussion of the resulted behavior of using OCO-3 XCO_2 only, OCO-2 XCO_2 only, and OCO-2&OCO-3 XCO_2 combined. Please see below sections for detailed comments and I would expect the manuscript to be published once comments and questions have been resolved.

Main comments/questions:

More information needed for the GCASv2. I understand the GCASv2 is an established model and described in at least two other published journal articles, but detailed information on the model setup, inversion methods, and error covariance metrics can be very helpful for readers of this manuscript to better understand the inversion system and results. Also see related comments in the Technical notes section.

Response: Thank you for this suggestion. We have added more information about GCASv2 in section 2.1 and the inversion settings in this study in section 3 in the revised manuscript.

In Section 2.1, the following paragraphs or sentences have been added (Pages 4-7, lines 107-166):

"……to implement the inversion of surface fluxes. MOZART-4 is an offline global chemical transport model developed in the National Center for Atmospheric Research (NCAR). It can be driven by essentially any meteorological data set and with any emissions inventory, so there is not a unique standard simulation (Emmons et al., 2010). We turned off all gas-phase, heterogeneous chemical reactions, aerosol and deposition processes in the MOZART4 model and added a corresponding number of CO_2 tracers according to the ensemble number in GCASv2, in order to allow the model to run more quickly. EnSRF assimilates observations in a sequential way, and obviates the need to perturb the observations. It shows good performance as long as the observation errors are uncorrelated (Houtekamer and Mitchell, 2001). GCASv2 is an upgrade from the GCAS (Zhang et al., 2015) that was established in 2015. The main upgrades include: 1) the addition of an assimilation module for satellite observations; 2) a change in the assimilation algorithm (i.e., EnSRF); 3) a change in the operational flow of the assimilation system; 4) the addition of a 'super-observation' scheme; 5) inversion of fluxes at the grid scale; and 6) an improvement in the localization scheme.

GCASv2 runs cyclically, with a two-step optimization strategy in each assimilation window (1 week). In the first step, the prior fluxes ($\boldsymbol{X_0^b}$) in each grid are independently perturbed with a random number (δ_i) drawn from a Gaussian distribution with mean of 0 and standard deviation of 1, and a scaling factor (λ) that represents the uncertainty of each prior flux (Eq. 1).

$$\boldsymbol{X_i^b} = \boldsymbol{X_0^b} + \lambda \times \delta_i \times \boldsymbol{X_0^b} \ \ , i = 1, 2, \dots, N \tag{1}$$

Then, the perturbed fluxes are put into the MOZART-4 model to simulate ensembles of CO_2 concentrations. The CO_2 profiles are sampled according to the locations and times of XCO_2 observations and converted to the simulated ensembles of XCO_2 ($XCO_{2,i}^m$) according to prior XCO_2 (XCO_2^a), prior XCO_2 profiles ($y_{a,j}$), pressure weighting function (h_j), and averaging kernel (a_j) of the XCO_2 retrievals (Eq. 2).

$$XCO_{2,i}^m = XCO_2^a + \sum_j h_j a_j (A(CO_{2,i}) - y_{a,j}) \tag{2}$$

Subsequently, the perturbed fluxes ($\boldsymbol{X_i^b}$), the simulated XCO_2 ensembles and the observed XCO_2 (y) are used in EnSRF to optimize the carbon fluxes ($\overline{\boldsymbol{X^a}}$) (Eqs. 3-5). The background error covariance matrix ($\boldsymbol{P^b}$) is calculated based on $\boldsymbol{X_i^b}$ according to Eq. (3), where $\overline{\boldsymbol{X}}^{\boldsymbol{b}}$ is the mean of $\boldsymbol{X_i^b}$. The posterior flux ($\overline{\boldsymbol{X^a}}$) is a correction to the prior flux using the bias between simulated and observed XCO_2 ($\mathbf{y} - \boldsymbol{H}\overline{\boldsymbol{X^b}}$) and the Kalman gain matrix ($\boldsymbol{K}$) (Eq. 4). And $\boldsymbol{K}$ is calculated according to Eq. (5), which is a function of model-data mismatch error covariance matrix ($\boldsymbol{R}$) and the background error covariance matrix.

$$\boldsymbol{P^b} = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{X_i^b} - \overline{\boldsymbol{X}}^{\boldsymbol{b}})\,(\boldsymbol{X_i^b} - \overline{\boldsymbol{X}}^{\boldsymbol{b}})^T \tag{3}$$

$$\overline{\boldsymbol{X^a}} = \overline{\boldsymbol{X^b}} + \mathbf{K}(\mathbf{y} - \boldsymbol{H}\overline{\boldsymbol{X^b}}) \tag{4}$$

$$\mathbf{K} = \boldsymbol{P^b H^T}(\boldsymbol{H P^b H^T} + \boldsymbol{R})^{-1} \tag{5}$$

In the second step, the optimized carbon fluxes are put into the MOZART-4 model to obtain the initial field of the next assimilation window. This scheme allows compensation of inversion results between neighboring windows and mass conservation between flux adjustments and concentration changes.

…… In this method, it first calculates the simulated XCO_2 corresponding to each observed XCO_2 based on the observation time and location, and then, it performs a retrieval error-weighted average for all the simulated and observed XCO_2 falling within the same model grid in the DA window, respectively.

There are inevitably spurious correlations in the EnKF method, to reduce the effect of spurious correlations, a two-layer localization scale was adopted in GCASv2, which is used to select which observations can be used for the flux analysis for each grid. The localization technique is based on the correlation coefficient between the simulated XCO_2 ensembles ($XCO_{2,i}^m$) in each observation location and the perturbed fluxes ($\boldsymbol{X_i^b}$) in current model grids and their distances. The observations will be accepted for assimilation if the distance is less than 500 km and the correlation coefficient is greater than 0; and if the distance is greater than or equal to 500 km and less than 3000 km and the correlation coefficient should be significant ($p<0.05$). Otherwise, the observations are not accepted. The reason for this scheme is that considering the atmospheric horizontal diffusion, we believe that there must be a correlation between the flux of one grid and the concentrations in its neighbouring grids, and therefore observations are accepted as long as this correlation coefficient is greater than zero. In contrast, at distant locations (>500 km), where the effect of atmospheric horizontal diffusion is essentially negligible, the relationship between source and receptor is mainly due to atmospheric transport, and in order to minimize spurious correlations we require that such correlations must be significant. More details of the system can be found in Jiang et al (2021)."

In Section 3 (see Page 11, lines 245-254),

"……and the FIRE and FOSSIL emissions are prescribed. According to Eq. (1), the prior NEE and OCN fluxes were perturbed using Eq. (6).

$$X_i^b = \lambda_{NEE} \times \delta_{i,NEE} \times X_{NEE}^b + \lambda_{ocn} \times \delta_{i,ocn} \times X_{OCN}^b + X_{Fire}^b + X_{Fossil}^b, i = 1, 2, ..., N \quad (6)$$

where X_{NEE}^b, X_{OCN}^b, X_{Fire}^b, and X_{Fossil}^b represent the prior fluxes of NEE, OCN, FIRE, and FOSSIL, respectively; δ_i is random perturbation samples, which is independent between grids; λ_{NEE} and λ_{ocn} are the scaling factors for prior NEE and OCN fluxes, which were set to be 6 and 10 in this study, respectively. As described above, the prior fluxes have a spatial resolution of 1° × 1°, for $\delta_{i,NEE}$ and $\delta_{i,ocn}$, we adopted a spatial resolution of 3° × 3°, and the outputs of the posterior fluxes have the same spatial resolution with the prior fluxes, that means in each 3° × 3° grid, the prior fluxes were adjusted with a same factor."

How are the posterior fluxes constrained when there's no observation data in GCASv2? For example, in the Exp_OCO3 at high latitudes, I would assume the posterior fluxes are less updated and would be similar to prior fluxes since no new information has been presented to the inversion system, but Figure 3 and Figure 6 seem to suggest the posterior fluxes changed substantially when compared to prior. More information on the EnSRF would be helpful for the readers to understand the inversion process.

Response: Thank you for this suggestion. Since the atmosphere is moving, a change in flux at a certain location can cause a change in concentration downwind, i.e., observations downwind can sense the flux change at that location, and thus we can use observations downwind to constrain the flux in that area. At high latitudes, although there are no observations of the OCO-3, observations downwind this region will be absorbed for assimilation by two-layer localization technique. The two-layer localization was employed to filter the observations used in the inversion, mainly to reduce the effect of spurious correlations. In the revised manuscript, we have added more information about the EnSRF and the two-layer localization technique, which has been detailed in the content of the previous response.

I'm curious about the authors' insights on why in general using OCO-2 XCO_2 alone and OCO-2&OCO-3 XCO_2 combined outperforms the experiment using OCO-3 only? Would there be any other reason except the spatial coverage and potential bias in OCO-3 XCO_2(line 263)?

Response: Thanks! We further analyzed the reasons for the poor assimilation of OCO-3 XCO_2 alone, and found that, in addition to the absence of observations in regions beyond 52° North and South latitudes, the varied observations timing and the cyclical variations in the observation data amount had an important influence on the inversion results. We first examined weekly changes in the data amount of OCO-3 using the re-grided data as described in Section 2.3, and found that there are very significant cyclical fluctuations in the data amount from OCO-3 (Figure S4a). For the observation time, all observations of OCO-2 were at 1:30 p.m. local time (LST), whereas that of OCO-3 were variable, with only about 14% of the observations near 13:30 p.m. LST and about 54% in the morning or after 4:00 p.m. LST (Figure S1). In order to quantify these effects, we added 3 additional inversion experiments, which were named as Exp_OCO2r, Exp_OCO3tc, and Exp_OCO2ts (Table S1). In Exp_OCO2r, only the OCO-2 XCO_2 retrievals located between 52°S and 52°N retrievals were assimilated, in Exp_OCO3tc, all the observation times of the OCO-3 XCO_2 retrievals were changed to 1.30 p.m. LST, and in Exp_OCO3ts, only OCO-3 data with observation times between 12 and 3 p.m. LST were assimilated. We find that the lack of data beyond 52° North and South latitudes is the main reason for the poor assimilation of OCO-3, and the observation time as well as the cyclical variations in the observation number also have an important effect on the results. In the revised manuscript, we have added two long paragraphs to discuss the issue.

We have added the following paragraphs in Section 4.5 in the revised manuscript (see Lines 463-506, Pages 21-23):

"Since OCO-3 has similar observation uncertainties of XCO_2 with OCO-2 (Taylor et al., 2023), the poor performance of assimilating OCO-3 XCO_2 retrievals (Exp_OCO3) may be related to that 1) OCO-3 lacks observations beyond 52° North and South latitudes (Figure 1a); 2) the observation time different from OCO-2; and 3) its spatial coverage between 52°S and 52°N. We first examined weekly changes in the data amount of OCO-3 using the re-grided data as described in Section 2.3, and found that there are very significant cyclical fluctuations in the data amount from OCO-3 (Figure S4a). Every 8 weeks or so, there is a trough in the data amount. There is a difference of about 5 times between the weeks with the highest and the lowest data amount, and in the weeks with least data amount, there were essentially no observations in the northern hemisphere (Figure S4b). This implies that the surface carbon fluxes are largely unconstrained in the Northern Hemisphere, especially at mid- to high-latitudes, during the weeks with low observational data, resulting in poorer assimilation performance than for OCO-2. For the observation time, all observations of OCO-2 were at 1:30 p.m. local time (LST), whereas that of OCO-3 were variable, with only about 14% of the observations near 13:30 p.m. LST and about 54% in the morning or after 4:00 p.m. LST (Figure S1). For reasons such as coarser model resolution, the global atmospheric chemical transport models generally simulate atmospheric concentrations better only in the afternoon, when boundary layer heights are at their highest and atmospheric mixing is at its best, so assimilating these observations in the morning and after 4 p.m. LST may result in poorer inversions due to the greater simulation bias of the atmospheric transport models at these times of day.

In order to quantify these effects, we added another 3 additional inversion experiments, which were named as Exp_OCO2r, Exp_OCO3tc, and Exp_OCO2ts (Table S1). In Exp_OCO2r, only the OCO-2 XCO_2 retrievals located between 52°S and 52°N retrievals were assimilated, in Exp_OCO3tc, all the observation times of the OCO-3 XCO_2 retrievals were changed to 1.30 p.m. LST, and in Exp_OCO3ts, only OCO-3 data with observation times between 12 and 3 p.m. LST were assimilated. When the OCO-2 data beyond 52° North and South latitudes were also removed (Exp_OCO2r), the NEE estimates, both globally and for individual regions, are close to those of the Exp_OCO3 experiment, especially in the high latitude region of Europe and boreal North America, the inverted NEEs are almost identical to those of the Exp_OCO3 experiment (Table S2 and S3), and the bias of a posteriori concentrations from observations at high latitudes is close to that of the

OCO-3 experiment (Figure S3). However, globally, compared to the OCO-3 experiment, the Exp_OCO2r experiment still has smaller the deviation between the global net flux and the observed annual growth rate (Table S2), and smaller the global mean bias of the posterior concentrations (Table S4). This suggests that the lack of observations of OCO-3 beyond 52° North and South latitudes does have a significant impact on the inversion results. In addition, it can also be noted that at mid-latitudes, the bias of Exp_OCO2r is also smaller than the OCO-3 experiment, which may be caused by the significant fluctuations in the data amount of OCO-3 (Figure S4). When we changed all the observation times of the OCO-3 XCO_2 retrievals to 1.30 p.m. LST (Exp_OCO3tc), although we are not actually able to do so, the inversion does show a significant improvement compared to Exp_OCO3. However, if we only select the data with observation time between 12:00 and 3:00 p.m. LST (Exp_OCO3ts), the deviation between the global net flux and the observed annual growth rate, and the mean biases of the posterior concentrations at most latitudes are larger than those of Exp_OCO3 (Table S2 and Figure S3), indicating a poorer performance than Exp_OCO3. The probably reason is that the data number of observations is substantially reduced at this time (Figure S2), which leads to a substantial weakening of the observational constraints on surface carbon fluxes (Figure S5). ”

General comments:

Line 113: How does GCASv2 handle parameters of the aggregated 'super-observation'? For example, if multiple OCO soundings has been aggregated into one 'super-observation', how does GCASv2 incorporate information such as pressure weighting function and averaging kernels from each individual soundings?

Response: Thank you for this suggestion. The 'super-observations' are generated by averaging all observations within an assimilation window for the same model grid. In this method, it first calculates the simulated XCO_2 corresponding to each observed XCO_2 based on the observation time and location, and then, it performs a retrieval error-weighted average for all the simulated and observed XCO_2 falling within the same model grid in the DA window, respectively. In the revised manuscript, we have added the following sentence (Lines 148-151, Page 6) to make it clear.

"……a single high-precision "super-observation". In this method, it first calculates the simulated XCO_2 corresponding to each observed XCO_2 based on the observation time and location, and then, it performs a retrieval error-weighted average for all the simulated and observed XCO_2 falling within the same model grid in the DA window, respectively."

Line 131: Can you justify the use of ocean glint? Ocean glint data is in general avoided in inversions due to potential high bias.

Response: Thank you for this suggestion. Indeed, in most of the previous studies that used OCO-2 XCO_2 to invert surface carbon fluxes, the OG data were not used. The reason is that the OG XCO_2 may have larger uncertainties, inversions assimilating OCO-2 OG retrievals produced unrealistic results of annual global ocean sinks (Peiro et al., 2022). In addition to its large uncertainties, we believe that another reason for the poor assimilation performance of OG is the relatively homogeneous distribution of XCO_2 on ocean, causing a large correlation of the model-data biases among different XCO_2 observations within a same region, which leads to observations at the same region having the same direction of adjustment for surface fluxes, and thus leads to a significant overestimated or underestimated ocean carbon sink. Because of this, some assimilation algorithms (e.g., EnSRF) can only achieve better assimilation results when the model-data biases between observations have relatively small correlation or are uncorrelated. Therefore, in this study, we set the OG data with larger uncertainties than the LNLG data, and re-grided it at a coarser spatial resolution of 5°×5°. The results show that under this scheme, the inverted ocean sink is reasonable, with value of -2.6 PgC yr^{-1} (Table 1). According to the reviewer 1's suggestion, we have added three additional inversion experiments in the revised manuscript, in which we use only land nadir and land glint (LN+LG) OCO-2 and OCO-3 retrievals for the inversion (Named as Exp_OCO3L, Exp_OCO2L and Exp_OCO3&2L, respectively). We compared the estimates of NEE and the evaluations against *in-situ* observations between the experiments with and without OG data, and found that assimilating OG data with our method can improve the inversions somewhat compared to removing OG.

We have added a paragraph in Section 4.5 in the revised manuscript (see Lines 439-462, Pages 20-21):

"In most of the previous studies that used OCO-2 XCO_2 to invert surface carbon fluxes, the OG data were not used (e.g., Peiro et al., 2022; Byrne et al., 2023), the reason is that the OG XCO_2 may have larger uncertainties, inversions assimilating OCO-2 OG retrievals produced unrealistic results of annual global ocean sinks (Peiro et al., 2022). In addition to its large uncertainties, we believe that another reason for the poor assimilation performance of OG is the relatively homogeneous distribution of XCO_2 on ocean, causing a large correlation of the model-data biases among different XCO_2 observations within a same region, which leads to observations at the same region having the same direction of adjustment for surface fluxes, and thus leads to a significant overestimated or underestimated of ocean carbon sink. Because of this, some assimilation algorithms (e.g., EnSRF) can only achieve better assimilation results when the model-data biases between observations have relatively small correlation or are uncorrelated. Therefore, in this study, we set the OG data with larger uncertainties than the LNLG data, and re-grided it at a coarser spatial resolution of 5°×5°. The results show that under this scheme, the inverted ocean sink is reasonable, with value of -2.6 PgC yr^{-1} (Table 1). In addition, in order to compare the scheme that we have adopted in this study with the previous scheme that do not assimilate the OG, we added three additional inversion experiments, in which only the LNLG data were assimilated (Table S1). It could be found that all the three inversion experiments without OG observations place smaller constraints on the ocean fluxes compared to the original experiments, with the posterior ocean fluxes remaining almost identical to the prior ocean fluxes. Correspondingly, the inverted global land sink as well as the sinks in most regions show a slight decrease (Tables S2 and S3). Evaluations in comparison with in-situ observations showed that there are some increases in the a posteriori concentration biases for all three experiments after removing OG. For example, for the experiments assimilating OCO-2 data, the mean bias increased from 0.02 to 0.14 ppm (Table S4). This suggests that assimilating OG data with our method can improve the inversions somewhat compared to removing OG."

Line 134: Please explain the regridding process. Does the regridding process refer to the 'super-observation' described in section 2.1? How did the XCO_2 values and parameters for each sounding been processed? Did you take the mean, or median, or other methods? And can you justify the method you used? How did you handle the outliers in the observations with one grid box? Also, for the 'super-observation', does it mean that for each model grid box, there's essentially only one observation being used by the model to constrain the posterior fluxes? If that's the case, why does the data amount (Figure 1) matter (except for the grids containing 0 OCO soundsing)?

Response: Thank you for this suggestion. The re-griding process was performed during the pre-processing of satellite data and does not involve the 'super-observation' process. The OCO observations are filtered using the parameter of XCO_2_quality_flag, which indicates the quality of the data. Only data with XCO_2_quality_flag equal 0 was selected. Then, the observations of LNLG were re-grided into 1°×1° grid cells, and those of OG were re-grided into 5°×5° using the arithmetic averaging method. The other variables like the column-averaging kernel and the retrieval error, which are provided along with the XCO_2 product, are also dealt with using the same method. This process is the same as Wang et al. (2019). For the 'super-observation', it mean that for each atmospheric transport model grid box, there's indeed only one observation being used by the model to constrain the posterior fluxes, but for each grid's flux, it is not only constrained by the observations of the grid it is on, because the atmosphere is moving and its downwind observations can all be used to constrain the flux of this grid, and in the system we use a two-layer localization scheme to select the surrounding and downwind observations that are used to constrain the flux of that grid. Therefore, the amount of observed data can have a significant impact on the inversion results. In the revised manuscript, we have further explained the 'super-observation' scheme (see Lines 148-151, Page 6) and also provided a detailed description of the localization technique (see Lines 152-166, Pages 6-7).

Line 209: Could you list out the annual CO_2 growth rates for 2020-2022 that you used to calculate the average growth rates?

Response: Thank you! When we conducted the inversion work, GCB2023 (i.e., Friedlingstein et al., 2023) had not yet been released, we used the 2020 and 2021 data (4.99 and 5.23 PgC/yr) from GCB2022, as well as the Annual Mean Global Carbon Dioxide Growth Rates (2.2 ppm) in 2022 reported by NOAA Global Monitoring Laboratory (https://gml.noaa.gov/ccgg/trends/gl_gr.html) by multi-by a factor of 2.124. The average atmospheric CO_2 growth rate is 4.96 PgC yr-1 for 2020-2022. We compared the results in GCB2022 and GCB2023 and found there are some differences in these values. In GCB2023, the CO_2 growth rates from 2020 to 2022 have been updated to 4.97016, 5.2038, and 4.63032 PgC/yr, with mean of 4.93 PgC/yr. Therefore, in the revised manuscript, we have updated this value to 4.93 PgC/yr.

Line 214: Why does the joint assimilation of OCO-2 and OCO-3 XCO_2 give the best performance on a global scale? One potential reason-spatial coverage of OCO-3 XCO_2 has been mentioned briefly in several places in the manuscript, but an in-depth discussion would be expected.

Response: Thank you for this suggestion. The OCO-3 satellite observations have a sufficient number of observations in the mid-latitude land region, while the OCO-2 satellite observations have a wide spatial coverage, even at high latitudes (Figure 1 in the original manuscript). Therefore, Exp_OCO3&2 assimilates sufficient observations in the mid-latitude region and observations in the high-latitude region, and has the advantages of OCO-2 and OCO-3 at the same time. At the same time, the joint assimilation of OCO-2 and OCO-3 XCO_2 also absorbs more observations than assimilating the OCO-2 or OCO-3 alone, which will also make the assimilation better. Assimilating OCO-3 XCO_2 alone has poor performance, the reasons are that, on the one hand, the fact that it is only available between 52°S and 52°N, which leads to a lack of observational constraints on the carbon sinks at high latitudes, and there are the large fluctuations in the amount of observational data, which leads to significant differences in observational constraints at mid-latitudes at different times; on the other hand, its varied observation time also affect the inversions, but even choosing afternoon observations does not improve the inversions because the amount of observed data drops significantly. Therefore, a better option for the future would be to jointly assimilate the OCO-2 XCO_2 data and the OCO-3 XCO_2 retrievals observed in the afternoon (12:00 to 16:00 LST). We have added a detailed discussion about this issue in Section 4.5 of the revised manuscript (see Lines 463-506, Pages 22-23).

Line 221 - 224: Is the word 'sinks' in line 22 a typo? Otherwise the sentence does not make sense – the listed locations seem to have positive NEE values suggesting being CO_2 sources.
Response: Thank you! Yes, it is a typo. We have changed 'sinks' to 'sources' (see Line 283, Page 12).

Line 236: I would suggest the authors avoid using 'peaks' when describing the negative values to clear confusion, or maybe specify the values when doing comparison. For example, the 'peaks' for ExpOCO2 and Exp_OCO3&2 are higher than the prior when rotation 90 degrees for Figure 3 (f) and (i), but the actually corresponding values at the 'peaks' are lower because they are CO_2 sinks and the NEE values are negative. Same for 'the lowest peak' in line 237.
Response: Thank you! We have revised that sentence (see Lines 296-299, Page 13) as follows: "The posterior and prior fluxes have a similar distribution trend along the latitude, with a significant peak of carbon sink near 60°N, and the strongest sinks of Exp_OCO2 and Exp_OCO3&2 are comparable, which are significantly stronger than the a priori, while Exp_OCO3 has the weakest peak of carbon sink and that is close to the a priori."

Line 251: Potential confusion – by the word 'lower' do you mean the NEE value is lower (strong sinks) or the NEE value is higher (weaker sinks)?
Response: Many thanks for this suggestion. We mean that in all regions except temperate N. America, northern Africa, temperate Asia, and Australia, Exp_OCO3 shows a weaker carbon sink than Exp_OCO2. We have corrected it in the revised manuscript (see Line 309, Page 13).

Line 301: Which experiment are those numbers from?
Response: Thank you for this suggestion. These numbers are calculated by averaging all the 3 inversion experiments. We have revised that sentence to make it clear (See line 362, page 17).

Table 2 and Figure 4: Is the information presented in Table 2 and Figure 4 largely duplicated? If so, authors may consider removing Figure 4 if additional paragraphs are needed.
Response: Thank you for this suggestion. Figure 4 in the original manuscript is actually a visualization of the data in Table 2, so there is indeed a duplication of content. In the revised manuscript, we have removed Figure 4.

Figure 6, Figure 3 and Line 358: For high latitude areas (> 60 degree N), why is the BIAS from Exp_OCO3 not consistent with prior fluxes? Given the fact that no OCO-3 observations available beyond 52 degree north, I would expect the posterior fluxes are very similar to prior fluxes in high latitude areas since no observation can be used to constrain and optimize prior emissions, yet both Figure 3 and Figure 6 showed substantial changes when comparing posterior to prior from Exp_OCO3. It's possible that fluxes in high latitude can be updated due to spatial covariance assumed in the inversion system, therefore more details on the GCASv2 is needed in Section 2.1.

Response: Thank you for this suggestion. Since the atmosphere is moving, a change in flux at a certain location can cause a change in concentration downwind, i.e., observations downwind can sense the flux change at that location, and thus we can use observations downwind to constrain the flux in that area. In this study, we use a localization scale of 3000 km, which means that observations within a 3000km radius of a grid can be used to constrain the fluxes in that grid as long as they meet the localization requirements as described in section 2.1 in the revised manuscript.

We have added more information about the two-layer localization scheme (see Lines 152-166, Pages 6-7) as follows:

"There are inevitably spurious correlations in the EnKF method, to reduce the effect of spurious correlations, a two-layers localization scale was adopted in GCASv2, which is used to select which observations can be used for the flux analysis for each grid. The localization technique is based on the correlation coefficient between the simulated XCO_2 ensembles ($XCO_{2,i}^{m}$) in each observation location and the perturbed fluxes ($\boldsymbol{X}_i^b$) in current model grids and their distances. The observations will be accepted for assimilation if the distance is less than 500 km and the correlation coefficient is greater than 0; if the distance is greater than or equal to 500 km and less than 3000 km and the correlation coefficient is significant ($p<0.05$), the observations will be accepted for assimilation. Otherwise, the observations are not accepted. The reason for this scheme is that considering the atmospheric horizontal diffusion, we believe that there must be a correlation between the flux of one grid and the concentrations in its neighbouring grids, and therefore observations are accepted as long as this correlation coefficient is greater than zero. In contrast, at distant locations (>500 km), where the effect of atmospheric horizontal diffusion is essentially negligible, the relationship between source and receptor is mainly due to atmospheric transport, and in order to minimize spurious correlations we require that such correlations must be significant."

Line 367: Could the bias exist prior? If there's no OCO-3 observation available in high latitudes, how can the OCO-3 observation introduce additional bias?

Response: Thank you for this suggestion. As the response in the previous comment, although the OCO-3 satellite has no observations at high latitudes, the observations downwind that area can be used to constrain the flux in that area. However, the assimilation of OCO-3 is much less effective compared to the OCO-2 satellite, which has observations in high latitudes, because only distant observations can be used in the Exp_OCO3 experiment.

Line 376: period '1'?

Response: Thank you for this suggestion. We mean the period from 1 August 2019 to 31 December 2022. We have corrected it in the revised manuscript (see Line 509, Page 23).

Reference:

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., et al.: Global Carbon Budget 2023, Earth Syst. Sci. Data, 15, 5301–5369, https://doi.org/10.5194/essd-15-5301-2023, 2023.

Peiro, H., Crowell, S., Schuh, A., Baker, D. F., O'Dell, C., Jacobson, A. R., Chevallier, F., Liu, J., Eldering, A., Crisp, D., Deng, F., Weir, B., Basu, S., Johnson, M. S., Philip, S., and Baker, I.: 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.

Wang, H., Jiang, F., Wang, J., Ju, W., and Chen, J. M.: Terrestrial ecosystem carbon flux estimated using GOSAT and OCO-2 XCO_2 retrievals, Atmos. Chem. Phys., 19, 12067–12082, https://doi.org/10.5194/acp-19-12067-2019, 2019.