Title: Monitoring and quantifying CO2 emissions of isolated power plants from space
MS No.: egusphere-2022-1490

Below we reply to the reviewer comments point by point. The reviewer comments are shown in *italic*, and corresponding modifications and citations of the manuscript are quoted.

Report #1

1. *>> (2) For all 50 cases, the difference between the background calculated by the 99 percentile and the background calculated by the 60 percentile ranges from 0.23 to 0.77 ppm. The standard deviation of the background calculated for the 9 percentile bins for each case ranges from 0.08 to 0.26 ppm. […]*

   In response to the sensitivity regarding the background information mentioned above, we have made changes to the manuscript in L280-286: "In this study, the background for the cross-sectional flux method is determined by fitting of Eq. (3), while the background for the Gaussian plume model method (GPM) is determined by the 90th percentile. The difference in background obtained by these two methods is very small (Fig. S11a), with a maximum difference of 0.86 ppm and a minimum of 0.004 ppm (Fig. S11b). Under the two background calculation methods, the GPM method has good consistency in the estimation results driven by three wind field (Fig. S11c-e). With the background computing by Eq. (3), the conclusion that estimated emissions have better accuracy using the WPBL is still valid (Fig. S12)." <<

   Comment: Thank you for addressing this comment and for the additional analysis. A bias in the XCO2 background ranging from 0.23 to 0.77 ppm using different percentiles is quite large given that a 0.5 ppm bias in the background translates to about 10% bias in the estimated emissions for a XCO2 enhancement of 5 ppm. I would therefore not use the term "very small" to describe the result.

   Furthermore, you picked the 90th percentile for your study. Can you provide an argument for this specific value?

Response: "very small" is referring to the difference between the two methods, which is on average rather small, but indeed the maximum in the difference is rather high. Thus, we adapted the text in L280-286: "The difference in background obtained by these two methods is on average small (Fig. S11a), but with a maximum difference of 0.86 ppm and a minimum of 0.004 ppm (Fig. S11b)."

In Fig. S2, we tested performance at 9 percentiles (60th, 65th, 70th, 75th, 80th, 85th, 90th, 95th, 99th). For each percentile, we calculated the coefficient of determination R2, root mean square error RMSE, and mean absolute error MAE between estimated and reported emissions. We found that for all cases, the error was smallest at the 90th percentile, so in the end we chose the 90th percentile. For the average emission and emission sum of each power plant, the error is smallest at the 75th percentile, so when we choose to calculate the background uncertainty, we choose 4 percentiles (75th, 80th, 85th, 90th). The relevant description is in L225-L228: "Here the background uncertainty $\varepsilon_{background}$ is computed from the spread in emission estimates using the 75th, 80th, 85th, 90th percentile to define the background values. This range of percentiles lead to the smallest difference with the reported emissions, as shown in Figure S2." We added the text in L280-281: "while the background for the Gaussian plume model method (GPM) is determined by the 90th percentile showing the lowest error for all cases in Figure S2a".

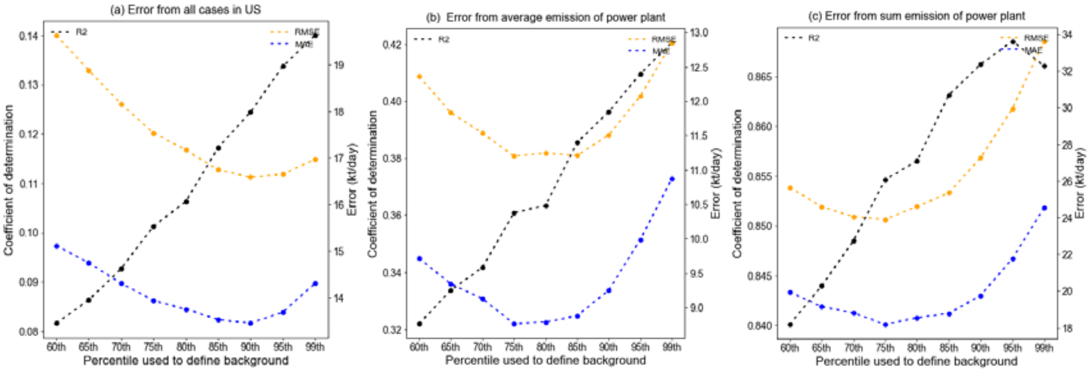

Figure S2. Different background results in different errors of estimated emissions and hourly reported emissions for isolated U.S. power plants. The errors of estimated emissions and hourly reported emissions by the GPM of all cases (a), the errors of the average value of emission estimation results of each power plant (b), and the errors of the sum of emission estimation results of each power plant (c).

2. *>> (3) In this study, we compared the effective wind computed from ERA5 and MERRA2 10 m wind speed with the wind at half the height of the PBL. […] Here, ERA-5 and operational forecast of ECMWF are very similar and assumed to perform in the same way. We didn't use the MERRA2 wind at half the height of PBL because of its lower resolution. <<*

   Comment: Thank you for addressing my comment. I agree that using half the height of the PBL should give good estimates of the effective wind speed, because it provides you a reasonable estimate of the mean wind speed within the PBL, where the CO2 plume should be well mixed during OCO-2 overpass in the early afternoon (see also Brunner et al. 2023, https://doi.org/10.5194/acp-23-2699-2023).

   However, I think you miss my main issue with your comparison. You compare three model products in your study: (1) ERA-5, (2) MERRA-2 and (3) ECMWF forecast. In addition, you compare two methods for computing the effective wind: (A) 1.4x 10-m winds and (B) winds at half the PBL height. As a reader, I like to know which model and method is most suitable for estimating power plant emissions. However, it is not possible to get this information from your analysis, because you only compare three options: 1A, 2A and 3B. I don't think it is enough to assume that ERA-5 and ECMWF forecast are similar. I would actually assume that ERA-5 performs better than ECMWF, because a reanalysis should be better than an operational forecast. I therefore think it is necessary to compare the three products using the winds at half the PBL height to obtain an objective result.

   I am not convinced that the 1.4-factor is a general value that can be applied for computing the effective wind speed for OCO-2 CO2 observations of power plants. In fact, Reuter et al. (2019) only used this factor "for convenience", while Hakkarainen et al. (2021) derives a scaling factor based on the surface pressure at the Matimba power plant, which (coincidently) was consistent with Varon et al. (2018). Varon et al. (2018) derive their factor specifically for CH4 plume observations with GHGsat instrument (50 m resolution, 1-5% instrument precision). The factor is directly linked to the detection limit and pixel size of GHGsat, because they only integrate over the detectable width of the plume. It is therefore not possible to generalize their results to OCO-2 CO2 observations with lower spatial resolution (2 km) and different detection limit. Anyway, I don't think it is necessary to have a detailed discussions on this topic, because you already conclude that using the 1.4-factor results in worse performance than using the *half-PBL value.*

(1) ERA-5, (2) MERRA-2 and (3) ECMWF forecast

(A) 1.4x 10-m winds and (B) winds at half the PBL height

Response: The comparison of option 1A and 2A is used to decide which meteorological field is performing better, ERA-5 or MERRA. This analysis clearly showed that MERRA is performing less, so that excluded the need for option 2A and 2B from the comparison. Subsequently, we conclude that the option (B) is better than option (A) and therefore 3A was no longer need checked.

ERA-5 is supposed to be better (for the past) than the operational ECMWF, but the differences in our case are small since we do not use any forecast of a few days ahead. Therefore, 1B and 3B are very similar and not compared, assuming that 1B is slightly better.

It would have been more thorough work if we had tested both 1B and 3B, but this would require a substantial amount of additional work and we prefer to leave this open for future research.

3. *>> (4) L316-317: "The total uncertainty is comparable to the uncertainty of power plant emissions in previous studies, which ranged from 3.42 to 19.2 (Nassar et al., 2017; 2022)" <<*
   Comment: Please add units.

Response: Thanks. We have added units in the revised manuscript, as follows:

"The total uncertainty is comparable to the uncertainty of power plant emissions in previous studies, which ranged from 3.42 to 19.2 kt/day (Nassar et al., 2017; 2022)."

4. *>> (4) "The uncertainty of wind speed is between 0.08 and 1.4 m s-1, and the uncertainty of background varies between 0.03 and 0.1 ppm (Table S1)." <<*
   Comment: The uncertainty in the background stated here is much smaller than the range from 0.23 to 0.77 ppm. What is the reason for the differences?

Response: The range from 0.23 to 0.77 ppm is for the 9 percentile bins (60th, 65th, 70th, 75th, 80th, 85th, 90th, 95th, 99th). The range from 0.03 and 0.1 ppm is for the selected percentile bins (75th, 80th, 85th, 90th) used to calculate uncertainty of background. More details are mentioned in the reply to comment 1.

5. *>> (5) The parameter A is one of the parameters to be fitted and not the line density. <<*
   Comment: This is wrong. The fitting parameter "A" in your Eq. 3 is already the line density (in ppm/m), i.e. the area under the Gaussian curve.
   If I understand your approach correctly, you first fit Eq. (3) and then use the fitting parameters (k and b) to compute the XCO2 enhancement by subtracting the background. You then use Eq. (S2) to compute the line density from the XCO2 enhancement, which you convert to g/m using Eq. (2)? If this is correct, I wonder if you approach has any advantage and how your line density actually differs from using the fitting parameter "A" directly.

Response: The line density unit is ppm*m. Thanks for the suggestion. We double check our line density calculation approach "the shaded area of Fig. R4" with the parameter A. We find the A is indeed same with the shaded area. Thus, we corrected the manuscript in L183-184: "A represents the line density which is same as the area under the fitted curve (Figure S1b) after removing the background."

OCO2-case-id	Fit_A	the shaded area (ppm*m)
1	9942.5	9942.5
2	47054.5	47054.5

3	7751.5	7751.5
4	14192.9	14192.9
5	10006.3	10006.3
6	13667.2	13667.2
7	15027.8	15027.8
8	10039.1	10039.1
9	11277.3	11277.3
10	14055.9	14055.9
11	17315.4	17315.4
12	13880.5	13880.5
13	27847.6	27847.6
14	9745.0	9745.0
15	33348.9	33348.9
16	4205.0	4205.0
17	7477.3	7477.3
18	5960.1	5960.1
19	16962.3	16962.3
20	7140.0	7140.0
21	15855.3	15855.3
22	39681.9	39681.9
23	28281.4	28281.4
24	10867.8	10867.8
25	4991.5	4991.5
26	12369.0	12369.0
27	12085.9	12085.9
28	28443.2	28443.2
29	4301.3	4301.3
30	5222.4	5222.4

6. *>> (9) We have added why WPBL provides better results in the revised manuscript L267-272: "[…]. The reason why the results using MERRA2 were worse (Fig. S4) is due to its low resolution, which cannot provide precise wind information for emission sources." <<*

   Comment: I think it would be important to add a reference to this statement.

Response: We added the MERRA-2 dataset reference mentioning the resolution in L271-272: "The results using MERRA2 were worse (Fig. S4) due to its low resolution (GMAO, 2015), which cannot provide precise wind information for emission sources."

7. *>> (16) We added explanations in the revised manuscript L378-382: "We found that the cross-sectional flux method has a larger variability than the GPM method. […] multiplied by the component of wind perpendicular to the orbit […]. <<*

   Comment: The explanation is still unclear to me. Do you mean that the limitation of the CFM is that you fit a symmetric Gaussian curve, but for a large angle between orbit and wind direction, the correct function would be an asymmetric Gaussian curve, which results in an additional source of uncertainty?

Response: Yes. When the angle between the orbit and the wind direction is large, the actual shape is asymmetric Gaussian. However, the resolution of OCO observations is not sufficient to fully fit asymmetric Gaussian curves. In addition, the use of wind speed components perpendicular to the track also resulted in estimated emission errors. Thus, we corrected the manuscript in L334-335: “This is because when the angle between the orbit and the wind direction is large, the actual cross-section shape is asymmetric Gaussian. But the resolution of OCO observations is not sufficient to fully fit asymmetric Gaussian curves.”