Point-by-point response to the reviewer's comments

We thank the reviewer Ray Nassar for the comments and suggestions provided, which have helped to improve the quality of the manuscript. We have taken into account all the suggested improvements:

Specific Points

Comment 1: Line 14: "possible thanks to" would better be rephrased as "made possible by"

Reply: We rephrased "possible thanks to" as suggested.

Comment 2: Line 60 and 61: capitalization of ENVISAT and TANSO is the advised, although TANSO-FTS is the complete name of the instrument.'

Reply: We changed "Tanso" to "TANSO-FTS" and capitalized ENVISAT.

Comment 3: 65: A Gaussian plume model does not account for eddies, however, it relies on the reasonable assumption that their effects are negligible for multi-kilometer spatial scales. It is recommended that the sentence is expanded to clarify this fact.

Reply: We expanded the sentence as suggested in L74 (of the revised version).

Comment 4: Line 116: "instantaneous hourly" would be more informative than just "hourly" to distinguish from an hourly average value.

Reply: We modified it accordingly.

Comment 5: Figure 1 caption "gross" should be "cross" or X.

Reply: We corrected the typo.

Comment 6: Line 156: Is there any justification of the requirement of less than 5 hours? Obviously a shorter offset in time is better, but are there any studies to quantify the effect that might justify this value? Both wind speed and direction could change significantly over a period of 5 hours, as discussed later around line 190.

Reply: The requirement of less than 5 hours is a generous *ad hoc* criterion. Despite the fact that both wind speed and direction could change over this time period, which would lead to different observed plume shapes for NO2 and CO2, small changes in the wind speed and direction in the time between overpasses do not play a significant role for the application of our method. This is because NO2 data is only used to define a potential plume that serves as a bounding box, i.e. it only constrains a spacial region to find the CO2 plume. In addition, this potential plume is the result of extending the detected NO2 plume, which increases the likelihood of the real CO2 plume being contained within the detected potential plume. Therefore, as long as there are no drastic changes in the wind speed and direction in the time between both overpasses, as is the case for the analysed scenes, the potential plume detected using NO2 will most likely contain the CO2 plume. The scene on 18 June 2021 (Fig. 7) is the only counter example. In this scene, the CO2 plume seems to not be fully contained within the potential plume (as mentioned in L636 of the revised version), which has however just a small effect for the emission estimation because the part of the CO2 that we miss is mostly farther than 35 km downwind of the source.

Thanks to this comment we have seen an inaccuracy in the manuscript. In L615 (of the initial version) we had mentioned, referring to the scene on 18 June 2021, that "the part of the CO2 plume that we miss is beyond the plume range, having no effect on the final result". However, in Fig. 7 we can appreciate that we miss part of the plume after about 31 km, leading to a small underestimation of the emissions in this case. We have corrected this in L637 of the revised version.

For a more systematic analysis of other scenes and targets, more filters need to be developed to automatically discard scenes where the detected plume only partially contains the CO2 emission plume due to changes in the wind speed and/or direction.

Comment 7: Line 208: This approach to account for swath bias is interesting and likely contributes to an improvement in emission estimates, however, should the swath numbering be "j = 1,2,... n", rather than only going up to n-1? Is it n-1 since the first swath has no offset, so j = 0,1,2 ... n-1, where s0 = 0?

Reply: The reason for the swath bias numbering being "j = 1,2,... n - 1" instead of "j = 1,2,... n", for a total of n swaths is that one swath has no offset because it was set as the reference. If we defined a n-th swath bias, s_n, Eq. 3 would lead to an under-determined linear regression without an

unique solution, i.e., the parameter a_0 could have any value and the different s_j would adjust accordingly. We can solve this indetermination by setting $s_n = 0$ as a reference, without loss of generality.

Comment 8: Line 277: 1-3 hours for the characteristic time used to determine the bottom-up value is consistent with the findings of Nassar et al. (2021, https://doi.org/10.1016/j.rse.2021.112579, e.g. Figure 1 and sec 2.5), which considered the plume extent, time since emissions to derive a time-weighted or 'dynamic' bottom-up value. This similar analysis is worth mentioning very briefly and citing.

Reply: We mentioned the similarity of both approaches and cited Nassar et al. (2021).

Comment 9: Section 2.3, uncertainty. Is there any uncertainty related to the observations? It was not entirely clear to me if this was indirectly included in the dispersion or sensitivity terms. The sensitivity term does account for uncertainty in the observations for background, but not necessarily the plume. Can the authors clarify?

Reply: We did not explicitly add an extra term related to the observations. We assumed that any biases in the plume observations are removed with the background subtraction. We clarified this in L314 (of the revised version) by adding the following comment: "We did not explicitly consider a XCO2 measurement error under the assumption that, at the relatively small spatial scales of the analysed scenes, any bias in the XCO2 data is corrected for when subtracting the background. Random errors in the XCO2 values are included in the dispersion uncertainty."

Comment 10: Line 559: "lead" should be "led"

Reply: Corrected.

Comment 11: Line 595: It is not surprising that the difference between applying quality filters and ignoring them reduced when observations near the Bełchatów lignite pit were excluded. The digital elevation model for OCO-3 v10 data does not account for recent anthropogenic effects on topography such as this, so biased XCO2 data will result through erroneous surface pressures. Although no DEM will be perfectly up to date with respect to anthropogenic effects on topography, the Copernicus DEM which

will be used in OCO-3 v11 data will reduce the problem and thus the difference between quality-filtering and not, will be reduced.

Reply: Indeed, it is not surprising that quality filtering reduces the observations near the lignite pit. As Christopher O'Dell suggested in their review to this manuscript, it might be warranted to filter out data points around the pit by hand because of the reason you state. Using a more accurate Digital Elevation Model (DEM) will presumably help to reduce the problem.