Response to Anonymous Referee #2

We thank Referee #2 for thoughtful comments and suggestions. We have addressed each of the comments below. Referee comments are in *red italics* and our responses are in Roman font.

> *In order to derive the emission ratios, using total column dry air-column mole fractions from satellite sounders one needs the know to what extent the observed variations in clean vs. urban air emissions are translated into the satellite signal. For NO2:CO ratios this was done by Lama et al. (2020) using CAMS simulated profiles onto which the total column averaging kernel was applied. Here the anomalies are divided by the surface averaging kernel values (I assume it corresponds with the value in the total column averaging kernel array that corresponds to the lowest altitude layer although this is not that clearly stated). Using these column averaging kernel values to look into aspects that pertain to a specific partial column, comes with its own set of uncertainties that need to be discussed. For instance the column averaging kernel values correspond to specific airmass layers with specific dimensions. These vertical dimensions could differ between sounders (and retrieved species), the Planetary Boundary Layer under consideration etc., and with that, biases could be induced into the analysis. This is particularly relevant for NO2, which features relatively low kernel values near the surface, but also CO which features an equally strong gradient near the surface (an additional figure showing typical total column averaging kernel profile shapes would be a useful addition to the paper).*

We use the column averaging kernel value at the surface pressure value of the measurement. This avoids issues with different retrieval grids and layers. We have now stated this explicitly. We have also included an example of the averaging kernel profiles at an enhancement in Appendix C.

> *Related to this, dividing the anomaly by the surface averaging kernel will only yield correct results if there is no a priori contribution to the anomaly (i.e. the a priori used in the satellite retrieval may not differ between what is considered background and urban air). This needs to be verified and clearly stated.*

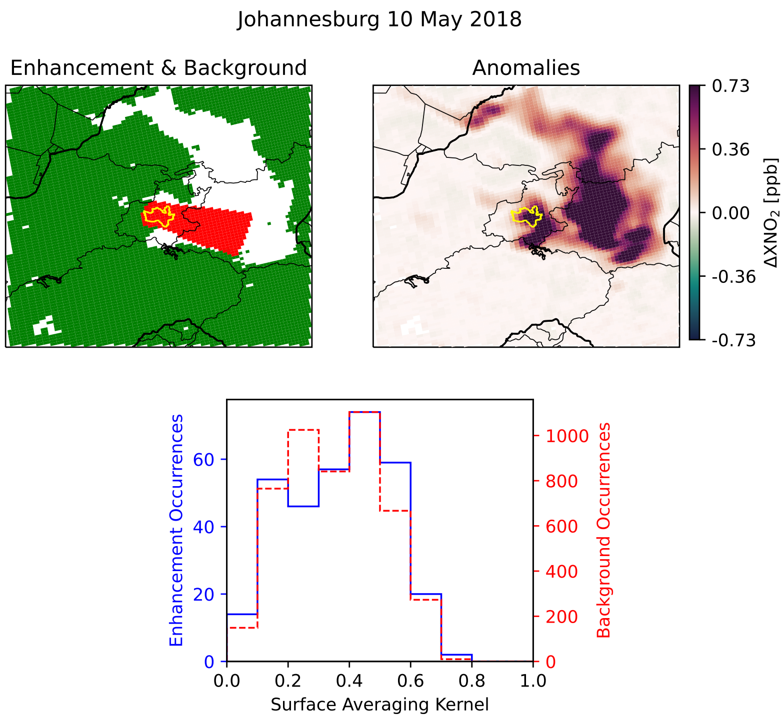
Thank you for pointing this out. The a priori CO_2 profiles for OCO-2 and OCO-3 depend only on dynamical latitude and time, and have no knowledge of urban vs rural locations. Therefore, the current method of dividing the measured anomaly by the surface pressure averaging kernel is valid for our CO_2 anomalies.

However, TROPOMI uses priors extracted from the TM5 chemical transport model on a 1° x 1° grid, and therefore the priors will contain urban enhancements. This requires a reformulation of the averaging kernel weighting, requiring a correction term:

$$\Delta c^t = \frac{\Delta \hat{c}}{a_0} - \frac{(1-a_0)\Delta c^a}{a_0}, \tag{1}$$

where Δc^t is the true anomaly, $\Delta \hat{c}$ is the measured anomaly, Δc^a is the prior anomaly ($\Delta c^a = c_u^a - c_b^a$, where c_u^a and c_b^a are the urban and background a priori columns, respectively), and a_0 is the surface pressure value of the column averaging kernel. The second term on the right hand side of the equation is required when $c_u^a - c_b^a \neq 0$, and reduces in magnitude as a_0 approaches 1. Equation 1, however, is only valid if the surface pressure averaging kernel value inside and outside the urban plume is the same (or similar enough). If the averaging kernel is tightly correlated with the atmospheric column, this equation is no longer correct, in general. The figure below provides an example of the similar distribution of surface averaging kernel values in the enhancement and background. The TROPOMI averaging kernels appear to be relatively insensitive to the plume itself, so we believe that Equation 1 is valid for our purposes. We have added the figure to the manuscript.

In the paper, we included the second term in our calculations and have rerun the analyses. The correction term was small for both NO_2 and CO, so the main results of the paper are unchanged, but the anomalies and enhancement ratios all changed slightly. We've also added an appendix to the paper containing the full derivation of Equation 1.



That said, we wanted to evaluate the magnitude of the error caused by excluding the extra correction term, so we performed a sensitivity analysis to quantify the magnitude of the bias caused by using the first term in Equation 1 over a range of surface column averaging kernel values and a range of a priori enhancements. We have not added this analysis to the paper, but it is here for completeness. In the analysis that follows, all the data are synthetic; there are no atmospheric measurements involved.

Figures 1–5 show in the top panel the true (modeled) enhancement in blue with square markers, the correction term (second term on the right hand side of Equation 1) in red with "+" markers, the approximation to the true enhancement (first term on the right hand side of Equation 1) in yellow lines with dots, and the simulated "measured" enhancement ($\Delta\hat{c}$) in purple lines with circles. All are plotted as a function of the surface pressure column averaging kernel value. The bottom panel shows the percent difference between the approximation to the true enhancement and the true enhancement (yellow), and the measured enhancement to the true enhancement (purple). Figure 1 shows that when the priors do not have an enhancement, the first term on the right hand side of Equation 1 is exact. Figure 2 shows the results when the a priori enhancement is equal to the measured enhancement. In this case, it would be more accurate to report the measured enhancement without a correction at all. The truth likely lies somewhere between these cases, where the retrieval makes a small adjustment to the a priori enhancement. If the a priori enhancement is between 0.5 and 1.5 times the measured enhancement, using the uncorrected measured enhancements incurs less error than using the first term on the right hand side of equation 1 alone (Figs. 3, 4). If the a priori enhancement is less than 0.5 of the measured enhancement, it is better to use the first term on the right hand side of equation 1 than the uncorrected enhancement (Fig. 5).

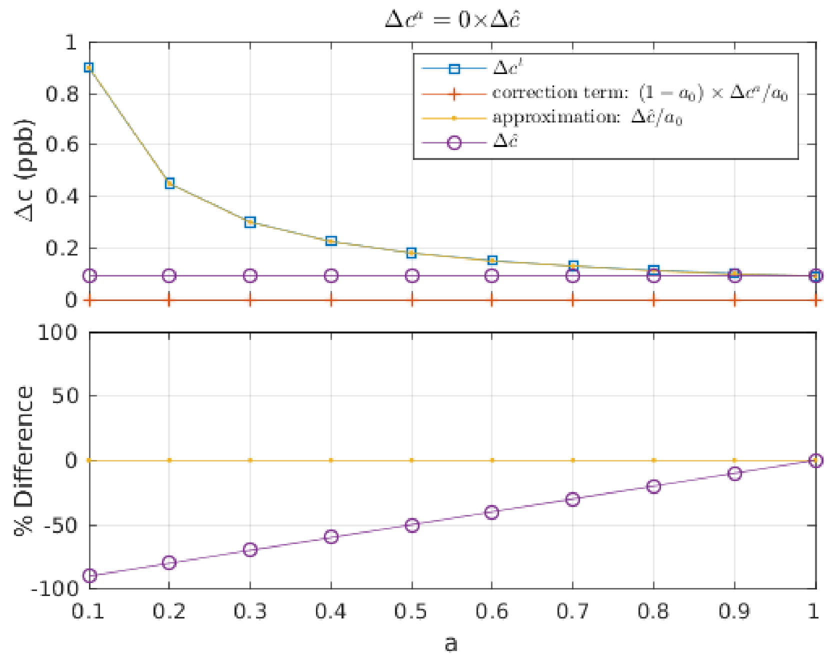


Figure 1: The impact of omitting the additional correction term (second term on the right hand side of Equation 1) on the corrected enhancement. In this case, the prior enhancement is set to 0, and therefore the first term on the right hand side of Equation 1 exactly reproduces the true enhancement.

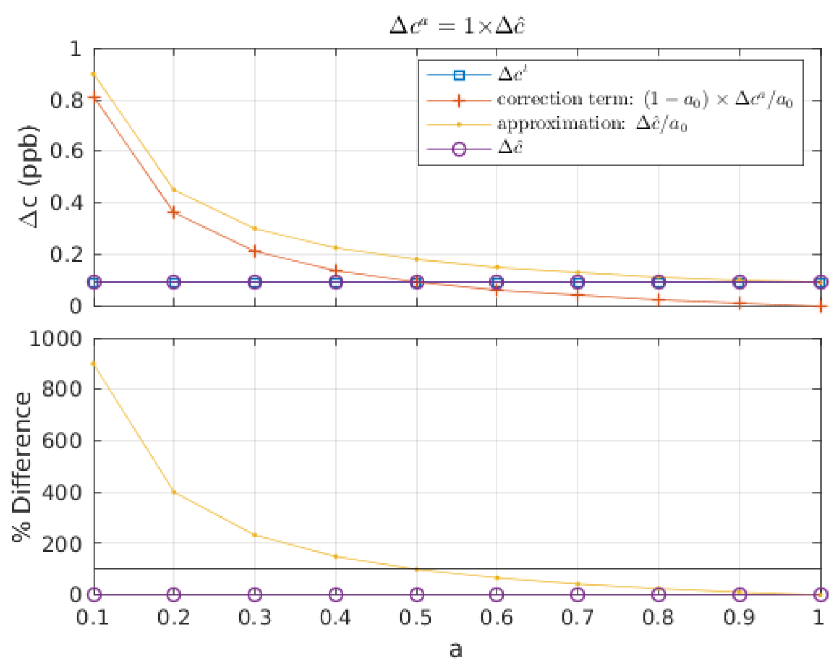


Figure 2: The impact of omitting the additional correction term (second term on the right hand side of Equation 1) on the corrected enhancement. In this case, the prior enhancement is set to be equal to the measured enhancement, and therefore the measured results are independent of the surface pressure averaging kernel value, and a better approximation to the truth than the first term on the right hand side of Equation 1.

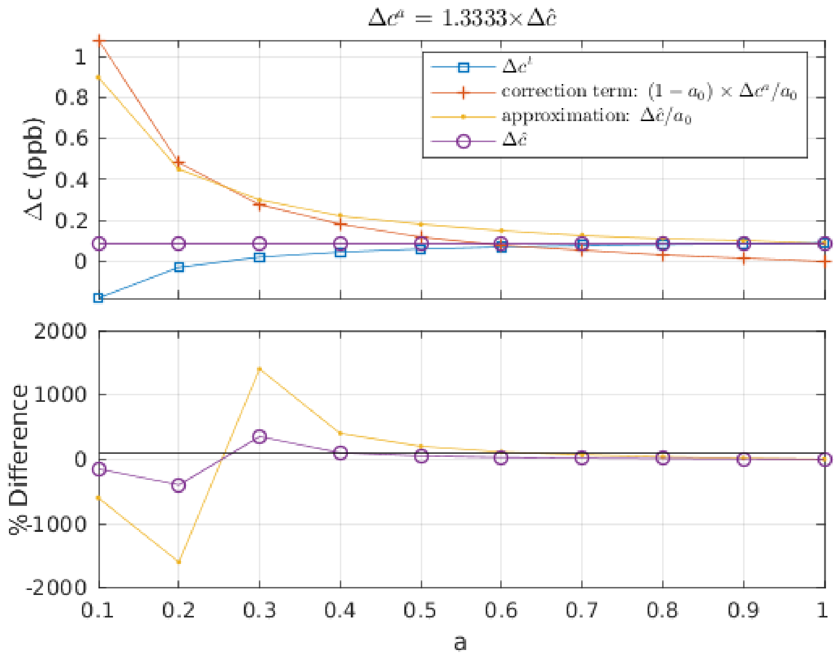


Figure 3: The impact of omitting the additional correction term (second term on the right hand side of Equation 1) on the corrected enhancement. In this case, the prior enhancement is set to be 33% larger than the measured enhancement. In this case, incorrectly adjusting the measured enhancement results in larger biases than using an uncorrected measured enhancement.

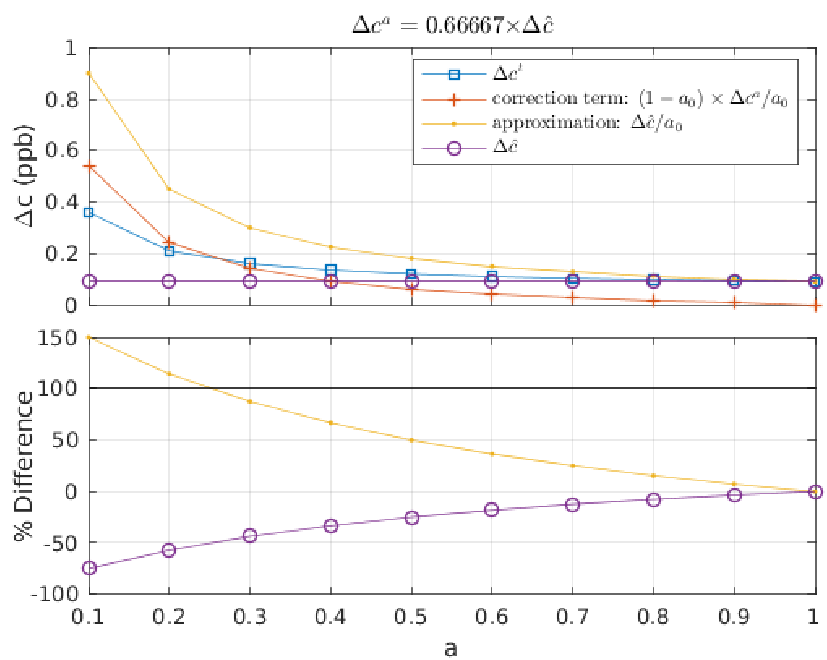


Figure 4: The impact of omitting the additional correction term (second term on the right hand side of Equation 1) on the corrected enhancement. In this case, the prior enhancement is set to be 33% smaller than the measured enhancement. In this case, incorrectly adjusting the measured enhancement results in larger biases than using an uncorrected measured enhancement.

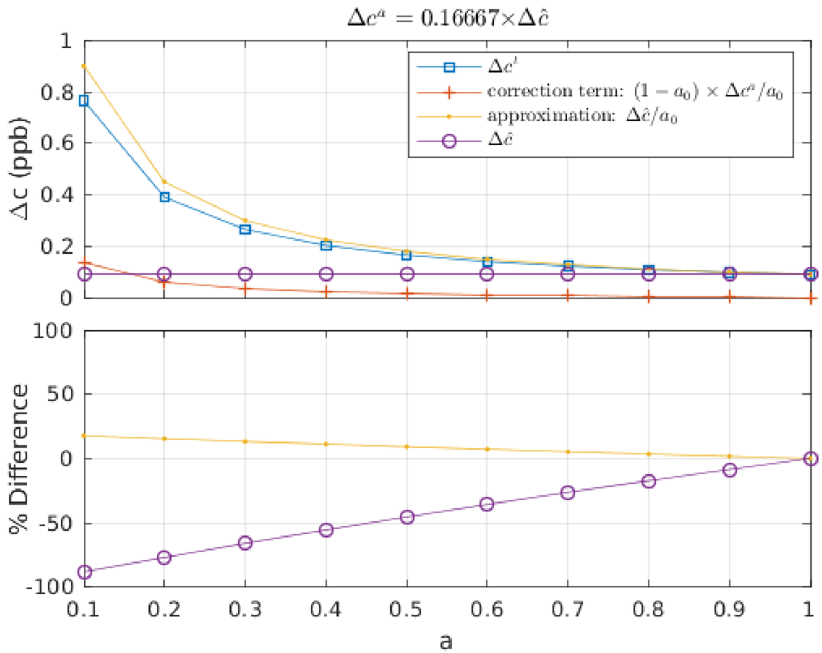
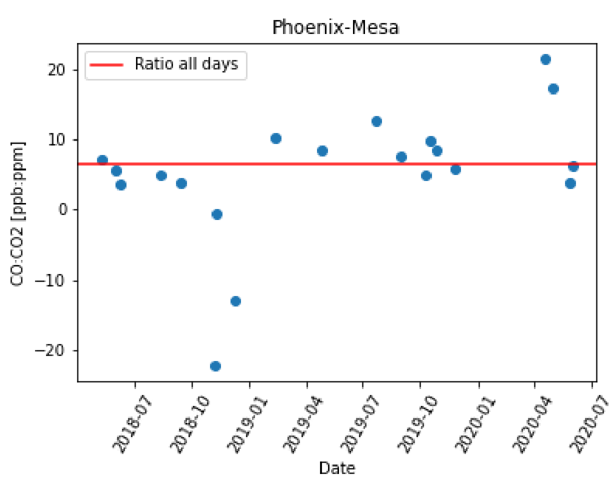


Figure 5: The impact of omitting the additional correction term (second term on the right hand side of Equation 1) on the corrected enhancement. In this case, the prior enhancement is set to be 83% smaller than the measured enhancement. In this case, incorrectly adjusting the measured enhancement results in smaller biases than using an uncorrected measured enhancement.

> *Secondly, some of the obtained emission ratios are derived from a very limited set of data. Foremost among these is Toronto with only one OCO-2/3 overpass. This begs the question to what extent this affects the overall uncertainty. Valid indicative information in this regard would be to perform the analysis for better sampled stations on an overpass per overpass basis and see what the obtained range of results would be and whether or not there is a temporal/seasonal aspect to the variability.*

We have performed the suggested analysis and include a figure below, showing that the overpasses are generally biased to summertime, but that we do not see much systematic seasonality in the enhancement ratios. That said, this representativeness error is not quantified in our analysis, and we have included a paragraph discussing this effect.



"A second possible source of error is the temporal representativeness of the satellite data used in this analysis. The overpasses that successfully pass our filtering criteria are biased toward sunnier conditions, and are most often collected in summertime, and some sites have very few overpasses (e.g., Toronto). If the enhancement ratios change seasonally, as expected, this type of analysis could cause a representativeness error, in which the comparisons between the measured enhancement ratios and the reported annual inventory ratios are systematically biased. Currently, the EDGAR and MACCity inventories, which provide CO and NO_2 emissions, do not report sub-annual emissions, so comparing to seasonal inventory ratios is not possible. With longer satellite time series providing more opportunities for wintertime enhancement ratios, we will be able to compute robust annual enhancement ratios to compare with the annual inventories."

Specific comments

> *Line 103: add references for each of these products*

Done.

> *Line 188: concerning 'points away': Is there a degree threshold?*

There is no explicit maximum angle threshold in our analysis, but the maximum angle between the wind and the center of the enhancement we use in this study is 66°; typically, the angle is less than 45°. In our analysis, the angle is permitted to be larger when the OCO-2 ground track is closer to the city. We have clarified in the text:

"When the boundary layer wind direction does not intersect the OCO-2 ground track the overpass is rejected, as the pollution plume from the city will not be captured."

> *Line 194: how are these corrections implemented? Could the misdirection of the plume also be related to the injection height of the emissions?*

It is possible that if the plumes are injected significantly above 50 m, (or if vertical mixing is much more vigorous than we expect), the 50 m winds we use would be inaccurate or inappropriate. We have added text to make it clear in the paper that we have assumed that plumes are injected around or below 50 m and that if this assumption is not true, the wind correction could partially correct for that:

"Errors in wind direction can be caused by the inability of the coarse model resolution to resolve local topography, or if the 50-m winds are not representative of the winds at the local plume height. The wind rotation we perform should at least partially correct for both these errors."

> *Line 228: Do you use certain fixed criteria to do this manual selection?*

We follow the manual selection methods used in Nassar et al. [2017, 2021], and visually identify a clear drop in XCO_2 to mark the end of the plume. An automated approach is worth considering for future work, but we have not yet identified a robust metric.

"a maximum downwind distance for the plume is determined manually, following Nassar et al. [2017;2021], to visually identify a drop in XCO_2, which limits the length of the plume to an area where significant enhancements are observed."

> *Line 291: assuming equal area?*

Correct. This has been clarified in the text.

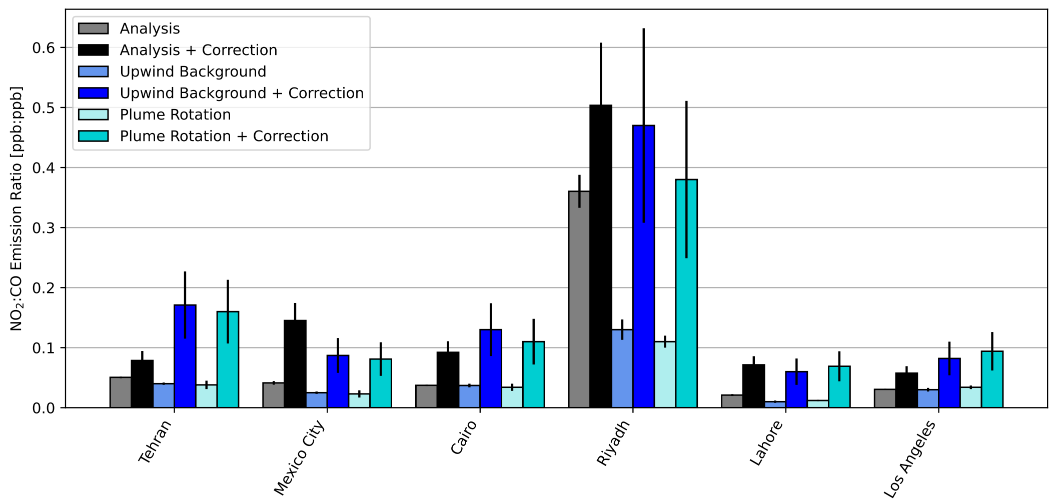
> *Line 295: 20% of the uncertainty itself, the initial emission ratio, the final emission ratio, the correction?*

We add a flat 20% in quadrature to the initial enhancement ratio uncertainty. This has been clarified in the text.

> *Line 357-364: It is stated that the differences between this study and Lama et al. prior to the lifetime correction falls within 25% for most sites. Could you discuss these differences in light of your (and Lama et al.) uncertainty estimates? An additional figure would also help the discussion. The authors also point to the difference in sampling rate (each overpass vs. a limited number of*

days in Lama et al.) but this aspect alone seems inconclusive to describe the particularly large discrepancies at some sites.

We have updated our comparisons with the Lama et al. paper. Our previous discussion was with reference to incorrect values in the appendix of Lama et al. (2020). We contacted the authors and they are in communication with the journal to fix this issue. We also added a figure to the text to show the differences more clearly. Our results show good agreement with the Lama results, within the uncertainties, of lifetime corrected NO_2:CO enhancement ratios at all cities except Tehran. An additional discussion has been added to the text.

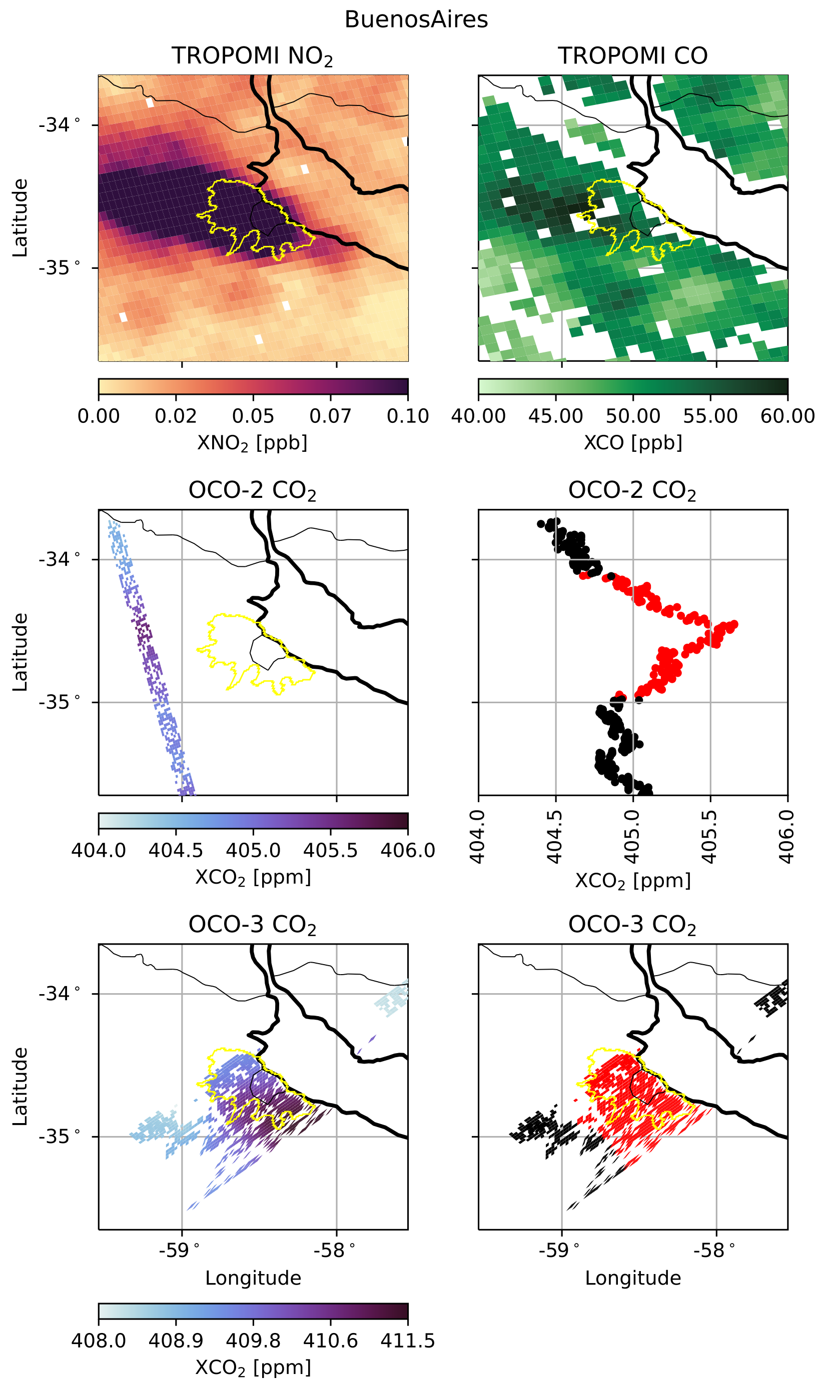


Line 448: Are we referring to lifetime corrected or uncorrected data? This is sometimes unclear in this section of the document. The agreement is sometimes different by several factors. 'Good' might therefore be nuanced.

We are focusing on the lifetime corrected results in this section, and we have added to the text to make this clear. The additional figure (see above) should help the reader understand what we mean by "good" agreement with the Lama et al. results.

Figure 1: It is unclear to me what the red area points to. It seems that this would be the downwind area that is considered for the analysis. But it also features some upwind areas?

In Figure 1 of the original manuscript, the data are along the OCO-2 ground track, and the OCO-2 track is downwind of the city. The red points are where the Gaussian plume intersects the OCO-2 track. We have clarified the caption to describe this more fully. We have also added a new figure (below) that that depicts the various satellite ground tracks that we hope will clarify where the red points come from.



Figures 6: Particularly Delhi features a correction that is larger than a factor 6. Could you elaborate why this particular city stands out?

We believe this is caused by particularly low wind conditions. The wind speeds used for the correction have been added to Table 3. We have added a comment about this in the text.

"Delhi has a particularly low wind speed (Table 3), which may cause the NO_2 lifetime correction to be overestimated."

References

Martínez-Alonso, S., Deeter, M., Worden, H., Borsdorff, T., Aben, I., Commane, R., Daube, B., Francis, G., George, M., Landgraf, J., Mao, D., McKain, K., and Wofsy, S.: 1.5 years of TROPOMI CO measurements: comparisons to MOPITT and ATom, Atmos. Meas. Tech., 13, 4841–4864, https://doi.org/10.5194/amt-13-4841-2020, 2020.

Borsdorff, T., aan de Brugh, J., Hu, H., Hasekamp, O., Sussmann, R., Rettinger, M., Hase, F., Gross, J., Schneider, M., Garcia, O., Stremme, W., Grutter, M., Feist, D. G., Arnold, S. G., De Mazière, M., Kumar Sha, M., Pollard, D. F., Kiel, M., Roehl, C., Wennberg, P. O., Toon, G. C., and Landgraf, J.: Mapping carbon monoxide pollution from space down to city scales with daily global coverage, Atmos. Meas. Tech., 11, 5507–5518, https://doi.org/10.5194/amt-11-5507-2018, 2018.

Laughner, J. L., Roche, S., Kiel, M., Toon, G. C., Wunch, D., Baier, B. C., Biraud, S., Chen, H., Kivi, R., Laemmel, T., McKain, K., Quéhé, P.-Y., Rousogenous, C., Stephens, B. B., Walker, K., and Wennberg, P. O.: A new algorithm to generate a priori trace gas profiles for the GGG2020 retrieval algorithm, Atmos. Meas. Tech. Discuss. [preprint], https://doi.org/10.5194/amt-2022-267, in review, 2022.