Review of

Comparison of OCO-2 target observations to MUCCnet - Is it possible to capture urban X_{CO2} gradients from space?

by Rissmann et al, 2022, submitted to AMT.

This paper compares small-spatial scale (<20 km) XCO2 gradients in OCO2 target-mode data with those from ground-based EM27 sensors located in Munich, Germany. Though nadir mode observations have analyzed some gradients, this is the first time multiple ground-based sensors have been used to evaluated OCO-2's ability to measure very small-scale XCO2 gradients, in particular from target-mode observations. This is important because it adds to the growing evidence that OCO-2 data does accurately capture small-scale XCO2 gradients, and therefore can be used to infer local-scale fluxes in CO2 from, e.g., urban areas and point sources (such as power plants).

This paper is very well-written and logically laid out. I have only some minor comments that should be addressed before publication. Once published, it will be an important addition to the literature on the topic of CO2 measurements from space.

My first major comment involves error bars. The error bars plotted on OCO2 biases (relative to MUCCnet), such as shown in Figures 7 & 8, appear to be standard deviation only. *IF* one were to assume that these errors were randomly distributed over the small area (<10 km) over which they are evaluated, we could estimate the standard error of the mean as sigma/sqrt(N), where N is the number of observations. However, it is well known that OCO-2 errors are NOT randomly distributed, over small or large areas (see e.g. Kulawik et al., 2019). It is likely that the local scale mean of OCO-2 includes systematic errors that cannot easily be evaluated. This point should clearly be made. Therefore, it is difficult to evaluate errors on mean OCO-2 values over the whole domain or the 3 sub-domains. Further, it is equally difficult to state them for the gradients. On this note, the error treatment on the gradients is both incorrect (as it assumes 1/sqrt(N) Gaussian averaging) and inconsistent with the stated error bars on XCO2 overpass means (which simply uses the standard deviation of the XCO2 values in a given domain). The current error bars listed on the gradients are unrealistically small (of order or less than 0.1 ppm). Therefore, please expand your discussion of errors to include these points, and be sure to treat errors consistent on the mean quantities and the gradient quantities.

Second comment (albeit minor) is how the data are presented on a map. Figure 2 shows this for the 12 dates analyzed. I suggest averaging the data onto a 0.5x0.5 km^2 grid, because as you show it now, it is large circles that overlap each other. Your plotting method emphasizes the noise in the data, rather than the spatial gradients. It further ignores that rather large horizontal extent of the OCO-2 FOVs, which are parallelograms and in some of these cases are rather wide! You can see an example of this in the figure below, which shows your plot of the Dec 18, 2020 case on the left, the corresponding NASA Worldview plot in the middle (obtainable at https://go.nasa.gov/38MQLRr), and my own home-made generated plot on the right. You can see the two on the right de-emphasize the noise and emphasize the gradients, and further do a better job illustrating the fact that OCO-2 target observations provide relatively complete spatial coverage due to the large degree of overlap of the various soundings in the image. The central image also illustrates the value of including the actual Aqua-MODIS data, at the expense of the spatial resolution of the surface imagery (unfortunately). But often there is cloud contamination in these Munich targets, so it is something to consider when plotting. You may wish to include something like the central image in your Figure 15, to really showcase the cloud context. And in general, please consider using one of these more realistic plotting methods which includes the actual parallelogram-shaped FOVs along with some kind of averaging.

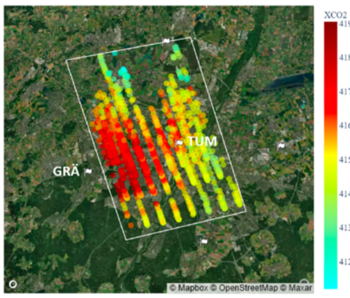
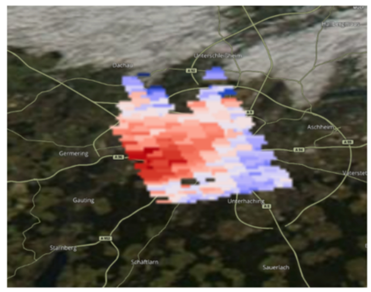
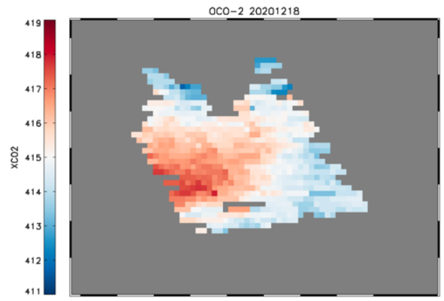


Fig 15(a) from this manuscript



NASA Worldview image, with colorbar ranging from 411 to 419 ppm. There is spatial averaging of the actual OCO-2 paraollelogram-shaped FOVs. The Aqua-MODIS true color image from the same orbit is shown as the background.



IDL-generated map by this reviewer, where the paraollelogram-shaped FOVs are averaged where they overlap, to a 0.006° x 0.006° (0.67x0.45 km^2) lat-lon grid.

Specific comments:

Section 3.2: The value of 6km radius around each EM27 seems rather arbitrary. Yes, it maximizes the circles while minimizing their overlap, but does it make sense as to the column of air the various MUCCnet sites are really sampling, as compared to OCO-2? For instance, using +/- 30 minutes overpass time for collocation, we could estimate that corresponds to a spatial extent of roughly +/- 15 km for a wind speed of 8 m/s, meaning that the EM27 signals will be much more “washed out” in the case of higher wind speeds, due to averaging all the EM27 values in that 1 hour time window. You should at least speak to this source of collocation error, and you may wish to mention that more sophisticated schemes (such as using a tighter time window) may reduce the collocation error.

Near line 150: Please state the spatial resolution of the anthropogenic emissions in TNO-GHG? It appears to be roughly 1x1 km^2, but if it is larger, than could also explain the Dec 18, 2020 discrepancy.

Near line 335: Please state the WRF-GHG wind speeds (roughly) in this case. If they are significantly larger than the 0.5-2 m/s wind speeds ERA5 shows, that could explain the ~factor of 10 discrepancy.

Near line 365: You ignore the factor of 10 discrepancy in the spatial XCO2 gradients between WRF-GHG and the observations. You should mention this as well, as the obs provide a potential way to improve whatever is going on in the model. I’m guessing perhaps there is something fundamentally wrong about how the model was set up and run, because I think WRF, at sufficiently high spatial resolution, should be able to duplicate the rough magnitude of observed spatial gradients.