State-wide California 2020 Carbon Dioxide Budget Estimated with OCO-2 and OCO-3 Satellite Data
By Johnson et al.

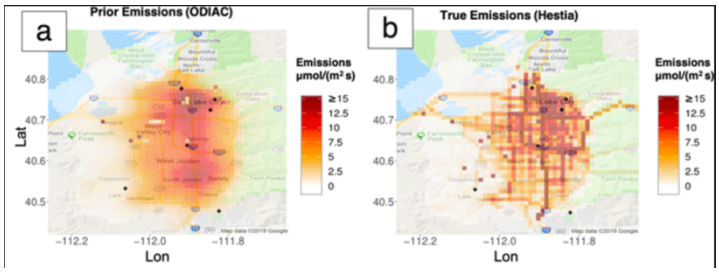
This work presented an inversion methodology capable of constraining statewide emissions across California. State of the art datasets and practices were used throughout and results demonstrated success. Total annual posterior estimates for 2020 (the "COVID" year) were in good agreement with prior estimates but certain sector-specific emissions (biospheric and wildfire emissions) demonstrated moderate disagreement. These disagreements led to interesting discussions about drought and fire characteristics during the year 2020. This work highlights the current and future potential of using space-based instruments to constrain CO2 emissions from not only CA, but other states as well. (Of course, global application is also possible, but some input datasets used are only U.S.-covering.) While the paper is overall scientifically sound and well written, there are a few aspects that I feel need some attention. I have arranged my comments such that major notes are listed first (which may require significant attention), followed by minor notes (which may be quickly addressed). For conversational reading/writing, I will refer to all authors as "you". Once these comments are adequately addressed, I recommend this manuscript for publication.

Major Notes

1. Apparently, California reduced CO2 emissions to 1990 levels. In **Lines 47-49**, you claimed that *"California was able to achieve this goal but in order to validate this, …, it is vital to have accurate estimates of past- and present-day greenhouse gas emissions."* This is an oddly worded statement since you claim that it was done but imply that it has yet to be validated. (How can you know it was done if it hasn't been validated?) My main complaint here is that the claim of CA reducing emissions to 1990 levels isn't backed up with a source. A curious reader may be interested in digging into how/when this happened. A source should be included here.

2. I am not familiar with all the studies listed in **Lines 75-82**, but this section seems to imply that the length scales in these studies are simply assigned and not estimated. I know that Roten et al., 2023 uses a variogram analysis on *each* SAM to determine the appropriate length scale. Thus, the length scale changes on a per-SAM basis. Do you still consider this to be "prescribed"?

3. What is the Vulcan emission inventory aggregated to? Table 1 lists its native resolution of 1km x 1km; however, the data presented in **Figure 1a** appear to be aggregated. (Perhaps it is merely an illusion?) Nonetheless, I find it odd that the range of flux is 0 to 4 umol/m2/s. This is the same range as NEE (-2 to 2 umol/m2/s)! Considering that similar emission inventories (ODIAC and Hestia) routinely present urban emissions greater that 15 umol/m2/s, the range presented in your figure is suspect. (See the provided figure from Kunik et al.,



2019). If your Vulcan input was in fact aggregated to a coarser resolution, this was not explicitly stated in the text. (Did I miss it somewhere?)

4. To my knowledge, the native temporal resolution of Vulcan is hourly; however, I assume V_{2015} is an annual average? (**Equation 2**). Did you use the hourly Vulcan data in any way? Additionally, I am a bit confused about C_{2020} and C_{2015}. Are these monthly as well, such that $(C_{2020})_i$ and $(C_{2015})_i$ where i is a particular month? Given how **Equation 2** is structured, it would make sense that they are annual estimates; however, C_{2020} is *also* used to create monthly scaling factors, R^{month}. The way **Lines 132-135** are worded induces some confusion about this. Obviously, if both C_{2020} and R^{month} are monthly data, this would lead to a redundancy. Can this be clarified?

5. How is ε determined in **Equation 4**? Where does its value(s) come from? I see that it is an error term but I don't think this is really explained within the text.

6. In **Line 338**, I think using "accurate" to describe the corrections applied to Vulcan via CARB is a bit of an overstatement. May I suggest something more conservative like "reasonable"? I say this because, while there is a clear "COVID dip" in the prior emissions during the spring and higher emissions elsewhere (summer, fall, and winter), a similar dip isn't as prominent in the posterior data (sd=0.5ppm). There is a similar dip in the Fall season. In particular, the prior and posterior appear to diverge in the fall, as the posterior decreases while the prior seems to increase. This *could* call into question whether the similar dip in the spring is due to COVID or not. (Perhaps the mechanism decreasing posterior emissions in the Fall is the same thing happening in the Spring?) Furthermore, since the uncertainties are large on posterior estimates, R^{month} values could be significantly increased/decreased while remaining within the posterior error bounds.

7. **Lines 357-359** and **365-367** are oddly worded.

   "Seasonal posterior NEE values displayed the largest uncertainty values of all source sectors in California and were on average ~95% of the seasonal posterior median emission value."

   "Seasonal posterior fire emission values displayed moderate to high uncertainty values and were on average ~80% of the seasonal posterior median emission values."

   Do the underlined statements apply to the "red" values or the "purple" values? I think my question is answered by the statement in **Lines 382-383** which says:

   "Observations as the 2σ standard deviation on this total flux is ~23% of the annual median posterior estimate."

   It seems the underlined statements above apply to the uncertainties (purple). Can you make this clearer?

Minor Updates

1. Obviously, OCO-2 and OCO-3 are the instruments being used in this study, but there are several other current and future missions that could be highlighted (GOSAT, GOSAT-2, TanSat, CO2M, etc.) (**Lines 60-61**)

2. In **Line 144**, "We et al." should be "Wu et al."

3. Just a suggestion for **Table 1**: include the total annual flux in the table but not in the bottom-line total. (Perhaps include an asterisk beside the value with a note that it isn't CA related?) Some readers may be curious as to what this value is. (This inclusion is certainly optional. Feel free to protest.)

4. The notation for R in **Equation 2** bugs me. For example, R^{month} *could* be seen as R^3 for the month of March. Obviously, the context of the equation suggests you wouldn't cube the value of R, but the notation could be clarified a bit.

5. "retrievals" should be "retrieval" in **Line 165**. No "s".

6. From **Line 204**, **Equation 7**: Is the Euclidean distance appropriate? Should it be the Haversine distance instead? Are the points spaced far enough apart for it to matter? *Is this the distance to aggregated OCO-2/3 soundings?*

7. In **Lines 214-215**, **Equation 9**, should this a Kronecker product? Certain Bayesian approaches distribute k_s onto k_t as a Kronecker product instead of an element-wise multiplication. (I'm not implying this is incorrect, I just want to double-check that the appropriate multiplication is, in fact, being used.)

8. Very minor comment: the word "using" appears twice in the same sentence and makes it a little awkward to read (**Lines 232-233**)

9. Another very minor comment: I noticed a pattern of abbreviating carbon dioxide as CO_2 and column carbon dioxide as XCO2. Is the inconsistency between the "2" subscript intentional?

10. It would be worth reminding readers that the prior mentioned in **Figure 4** is the *modified* Vulcan product. This could be done in the figure and/or in the caption. In **Lines 411-413**, you mention comparisons to CARB. It may be good to include a CARB estimate in **Figure 4** as well.

11. Can corresponding uncertainties be added to the values in **Table 2**?