Atmospheric
Chemistry
and Physics
Discussions

Interactive comment on "Linking global terrestrial CO$_2$ fluxes and environmental drivers using OCO-2 and a geostatistical inverse model" *by* Zichong Chen et al.

Julia Marshall (Referee)

marshall@bgc-jena.mpg.de

Received and published: 27 July 2020

At first glance it seems that the results of this study make sense, and are consistent with our general understanding of what drives carbon fluxes, with uptake at higher latitudes being mostly radiation-limited while in the tropics there are more complex temperature-precipitation interactions. So far, so good. The paper is well written and clearly structured, making it easy to read.

To the careful reader it soon becomes clear that something is going wrong, however, and the limited "validation" and comparison to other results from the literature are insufficient to explain these problems away. While the geostatistical approach is com-

mendable in that it allows more flexibility in the structure of the prior fluxes, such that perhaps unexpected signals may emerge, it also seems to allow for rather unphysical results, as in this case. Given the fact that the ocean fluxes (a net sink of more than 2 PgC/year) were rejected by the Bayesian Information Criterion (BIC) while the net land fluxes are more or less consistent with other studies, it seems impossible that the global atmospheric growth rate can be matched. It just does not add up.

This should be obvious when performing validation, but the very little testing of the posterior fluxes, limited to a handful of aircraft measurements far from coasts on a scatter plot averaged (monthly?) by height, hidden in the supplement, makes it hard to tell. The paper states that aircraft profiles near coasts were not used because the coarse model resolution made it hard to represent these data well, but I wonder if the complete absence of ocean fluxes may have also played a role here?

Since none of the in-situ sites were used for constraining the fluxes (which seems an odd choice, even if only for comparison's sake), it would be instructive to plot the concentrations resulting from the posterior fluxes at a few sites to see if the curves drift apart over the year as a result of the missing sink. While this might not look too bad in a simulation of only one year, this would soon result in wildly divergent curves. But perhaps over a longer simulation the BIC would then choose to select the ocean fluxes. Still, the decision to blindly allow the model to return what we know is incorrect makes it hard to trust the interpretation of the results. Perhaps Takahashi was not the best ocean prior in this case, especially for an El Niño year, and this played a role: this could be an area for more analysis.

The comparison to other model output was largely limited to the OCO-2 model inter-comparison study of Crowell et al. (2019), without following the considerable effort they put into validation or consideration of in-situ measurements. Looking at TCCON sites is an obvious choice, as is the extension to additional aircraft measurements, such as AToM, which are available for at least a couple months of 2016. But comparing your (unclosed) budget to the land biosphere budget of other (mass-conserving) studies is

intrinsically misleading. (I am not as surprised that BIC did not pick out the GFED emissions, as these are a few orders of magnitude smaller and are easily swallowed up in the biosphere signal.)

As the carbon budget presented in this study does not seem to add up in a basic back-of-the-envelope way, and the validation presented was not sufficient to identify that, I cannot recommend its publication without substantial revisions.

Other major comments:

L10 & L204-205: While the difference in wording is subtle, I think the abstract over-states what the meteorological variables explain. Do they really describe 90% of the variability in the fluxes (as seen through OCO-2 observations)? This sort of implies that OCO-2 can "see" fluxes, which isn't true of course. The latter explanation that the deterministic model accounts for XX% of the variance in the estimated fluxes seems more accurate. As you're only treating fluxes on a daily time scale, you're definitely not describing 90% of the variability in the fluxes themselves.

Figure 3 and discussion around L235: This is actually quite interesting! I would be interested in seeing some more analysis of this point. It was also not entirely clear to me what was correlated (and how) in Figure 3. The meteorological variables have been "passed through an atmospheric [transport] model": were they then sampled as column-averaged variables, as OCO-2 views the atmosphere? Were the same averaging kernels applied? It also says that this is the correlation "within different global biomes". Were these columns averaged across space then, and the correlation taken in time? Or is this a spatial correlation coefficient between the column-averaged maps for a given time? I feel like there is an intriguing result here, but I don't fully understand what you've done.

L238 & 239: How can you be sure that this collinearity is playing a bigger role than retrieval or model errors? Would the latter two effects not also limit the model selection?

L244 & L260: These statements seem to contradict each other. The first says that the negative beta values for PAR mean that an increase in PAR leads to a decrease in NEE and an increase in uptake. The latter says that the negative beta value for scaled temperature means that an increase in temperature leads to reduced uptake. How can these both be true? This is fundamental to the conclusions drawn.

L257-258: While cloudiness is correlated with clouds and rainfall, it's also correlated with the presence or absence of satellite measurements. What impact might this have on your results?

L302: I'm actually surprised Australia matches as well as it does, as you've had to fold the Southern Ocean sink into the Southern Hemisphere land fluxes somehow.

Minor comments:

In several places "in year 2016" should be replaced with "in the year 2016".

L42: refer -> referred

L46: levels -> level

L48: "At" should not be capitalized.

L55: region -> regional

L59 (and elsewhere): the ñ is in italics throughout.

L62: modeling -> model

L73: Here you mention that you are optimizing daily fluxes. Does this mean that the diurnal cycle is completely ignored?

L95: average -> averaged

L114 & 176: space before italic "p"

2020.

C5