Manuscript: Functional ANOVA for Carbon Flux Estimates from Remote Sensing Data

J. Hobbs, M. Katzfuss, H. Nguyen, V. Yadav, J. Liu

December 2023

1 Editorial Response

The topic editor report included some remarks that we address below

- **The manuscript briefly mentions “”functional analysis of variance (ANOVA)” for identifying important flux anomalies, but lacks a detailed explanation of the spatial-temporal statistical model’s construction and application. It is recommended that the authors provide a clearer methodological framework for functional ANOVA, especially in the context of its operationalization within spatial-temporal statistical models for estimating common flux signals.**

  We appreciate the reviewer’s request for additional clarification on the functional ANOVA methodology. The statistical model’s construction is outlined in detail in Section 3 of the manuscript with additional exposition in the supplement. Further, the implementation and application are discussed in subsequent sections. However, we agree that other sections of the manuscript and abstract could set the stage and reiterate the key methodological aspects more effectively. Therefore, the revised manuscript includes an additional sentence in the abstract that elaborates on the method. Additional clarifying statements are included in the Introduction and Conclusions.

- **The distinction between Level 2 and Level 3 products and their specific impact on flux estimation should be more thoroughly explained. It would be beneficial if the authors could conduct additional analyses to elucidate the impact of data product levels on the inversion results, thus strengthening the manuscript’s argument.**

  In the revised manuscript, we have clarified the discussion of the product levels in the Introduction. As noted in Section 2.1, none of the inversions studied actually ingest Level 2 products at their native resolution. Thus, the comparisons performed in our study, particularly for the fused CO_2 experiment, are comparisons of different approaches to constructing Level 3-type products.

- **While the manuscript acknowledges the influence of multiple sources of uncertainty on the flux solution, it could provide a more explicit description of how these uncertainties are integrated into the spatial-temporal statistical method and how the a posteriori uncertainty is represented. Elaborating on how different sources of uncertainty are managed within the method is recommended.**

  In this work, the partitioning of multiple sources of variability is achieved by defining different factors in the functional ANOVA implementation. Their relative importance is manifested, in part, through the Gaussian process (GP) standard deviations, as discussed in the Results section. We have added a sentence there to this effect. The Introduction notes that the flux a posteriori uncertainty is not always available, and this is the case for OCO-2 MIP products analyzed. We have added this note in Section 2 of the revised manuscript.

- **The manuscript delves into the influence of various factors on global flux estimation, including atmospheric transport modeling and data sources. However, it does not sufficiently discuss how these factors pertain to specific geographic regions and time scales. It is advisable to conduct more detailed regional and temporal analyses of these influencing factors to comprehensively understand their impacts.**

  Indeed, the impact of these sources of variability on the flux estimates can change in space and time. For this reason, we have studied their impacts through application of the functional ANOVA over three distinct continents, as discussed in Sections 3 and 4. While this is not a comprehensive set of regional analyses, it does span a range of geophysical conditions (tropics, mid-latitudes) and in situ data availability. For temporal coverage, the analysis is constrained in two aspects. First, the flux estimates are provided at monthly resolution; second, the OCO-2 record spans the recent decade. As the satellite CO_2 record continues, trends and interannual variability can be studied more thoroughly. In response to this comment, we have included the mutli-continent analysis in the Introduction and added some suggested future work in the Conclusions.

2 Referee Report 3

Additional comments from Referee Report 3 are addressed here

- **Modeling the "deviation" $(y - y^{(0)})$ is well motivated, but it is unclear from the paper what $y^{(0)}$ (prior flux) is specifically and how it's estimated/evaluated. Specifically writing this out in Sections 4.1 and 4.2 (or even in the supplement) for the data examined would be helpful.**

  The construction of prior fluxes for inversion systems is an art in its own right. Particularly for the OCO-2 MIP, each inversion system has its own prior flux fields, but these come from a relatively small collection of input sources. The prior fluxes attempt to capture climatological behavior of the biosphere, in combination with fossil fuel emissions and other known perturbations, particularly fires. We have a brief discussion on this in Section 3 of the revised manuscript, and the flux prior maps are available in the supplement.

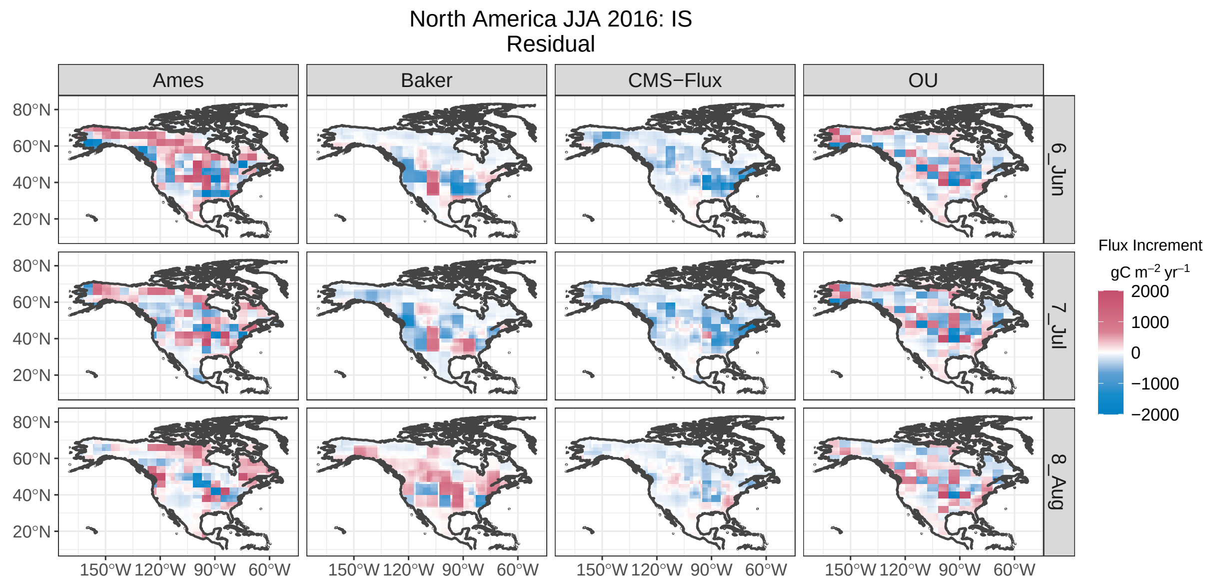
- **Because the error (epsilon) values are so large, I'd like to see a plot of these. Perhaps they could be made available in the supplement.**

  This is an interesting suggestion, and the functional ANOVA residuals can be assembled from the MCMC results. We include here an example (Figure EC.1) from the OCO-2 MIP analysis for North America. At this point, we do not think this additional exposition is warranted to include these results in the supplement.

- **I'd like more interpretation of the results and their implications.**

  - **An example of a place where more CO-2 and OCO-2 context would be useful is with Figure 10. Page 17 discusses that the estimated model effects for the mean flux increment changes sign across the equator or a "north-south contrast." What are the implications of this for OCO-2? Why does this matter for the MIP? Is this a surprising or expected result? In other words, how should scientists use this new knowledge/understanding?**

    This is an important point, and we have added further discussion on these differences in the revised manuscript, mentioning that north-south transport among the inversion systems could play a role. We also note that the density of OCO-2 observations changes from north to south across Africa.



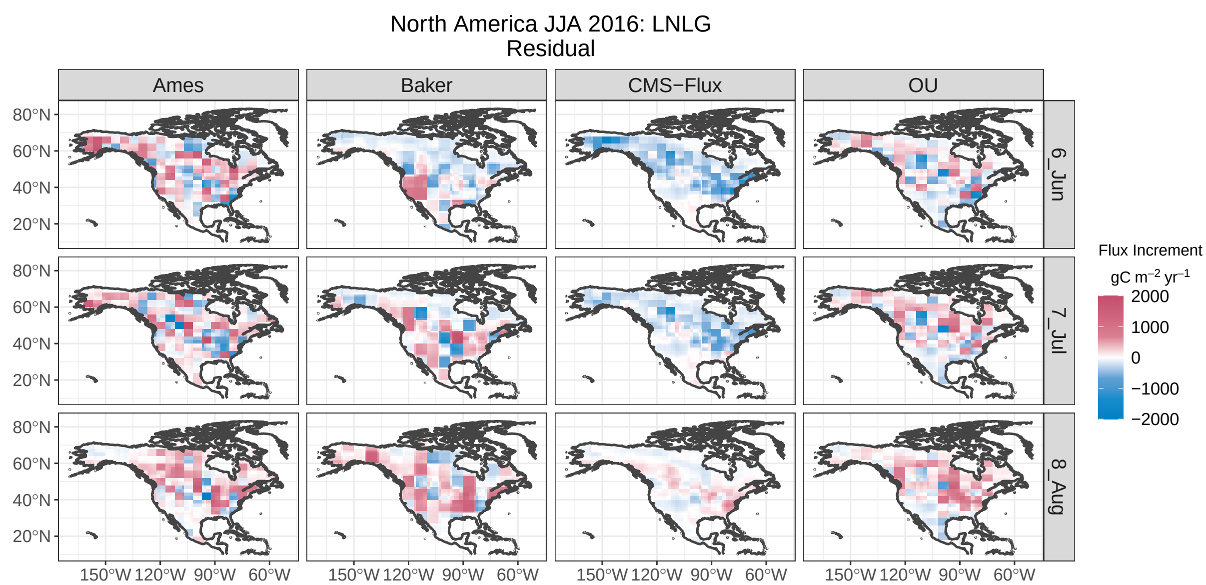


Figure EC.1: Functional ANOVA estimated residuals for the MIP example over North America. Units are gC m^{-2} yr^{-1}

  - **An example of cases where more statistical context would be useful is in regards to Figure 11. Figure 11 provides credible intervals and shows that most of the intervals include 0. As noted on page 18, places where the credible inferences do not include 0 are places where there are few/no in situ observation sites. Does it make sense to even make inferences here? How much are inferences at these locations biasing the results at other locations?**

    We appreciate this remark as well. In this instance, inversion systems' representation of atmospheric transport could be a factor yet again. The sparse in situ observations combine with the assumed atmospheric transport to yield the flux estimates. In fact, the reviewer's remark is an alternative interpretation of the data source effect; areas with sparse in situ coverage can have a relative "bias" in flux space versus the satellite inversions. We have added further interpretation in the revised manuscript.

- **Finally, the code is a bit overwhelming to try to reproduce the results. The "readme" files are helpful but I recommend having a "main" file where each relevant file is sourced in order to reproduce main results (even if it comes with a computation time warning).**

  We agree that there is a learning curve for the code repository. However, a driving script is not practical given the longer runtime for the MCMC sampler steps. With that said, we will update the examples' README to indicate which steps/scripts can be skipped over if a user simply wants to make use of the final MCMC samples that are available on Zenodo.