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

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

August 2023

1 Responses to Reviewer 1 (RC1)

1.1 Major comments

- **I'm concerned that one of the assumptions about the error distribution, namely that the realisations are independent, may not be valid. The realisations are taken as successive months within a three-month block. In my experience, geophysical fields and their associated error distributions are typically correlated in space and time. The authors need to address this issue directly, perhaps demonstrating numerically that the residual fields cannot be distinguished from random draws, or arguing persuasively why this should be overlooked for the sake of illustrative example. One thing to consider here, and probably should be raised in the revision, was whether the inversions were conducted as a "long" run (i.e., at least spanning all of the three months considered), or whether they were run separately for the individual months in question.**

  This is indeed an important point regarding the functional ANOVA model used in the first version of the manuscript. As the reviewer notes, all inversions are in fact performed as "long runs", considering the entire time record, so temporal correlation is possible. Upon receiving similar comments from both reviewers, we revisited the methodology and have modified the functional ANOVA model to allow for temporal correlation in the error terms $\epsilon_{ijk}(\mathbf{s})$. The modified statistical model is outlined in the revised manuscript in Section 3.1. The resulting estimates for the temporal correlation parameters are discussed with the other results in Section 4. The resulting estimates turn out to be relatively small but somewhat important for the overall inference, and the magnitude of the correlation is an interesting aspect of the analysis in its own right.

  The updated supplement includes some additional details on the specification of the temporal correlation structure and its integration into the MCMC algorithm. With the manuscript revision, we include a new release of the software repository that includes the spatio-temporal model for the error process.

- **It is unclear what is going on with the ocean areas. The distribution of fluxes is very different for land areas or oceans. Similarly, the distribution of fluxes depends very much on the type of vegetation cover. Consider for a moment differences between tropical rainforest vs savannah vs deserts vs high-latitude boreal forests. Are we assuming a uniform standard deviation? This should be discussed.**

  The analysis does not include fluxes over the oceans, and we provide additional detail on this in the revised manuscript and address in detail in the response to the next comment below.

  The additional remark on land/vegetation cover is an important comment. As currently implemented, the functional ANOVA Gaussian process models assume a constant variance for each of the components, including the error process. This combination does offer substantial flexibility to fit the

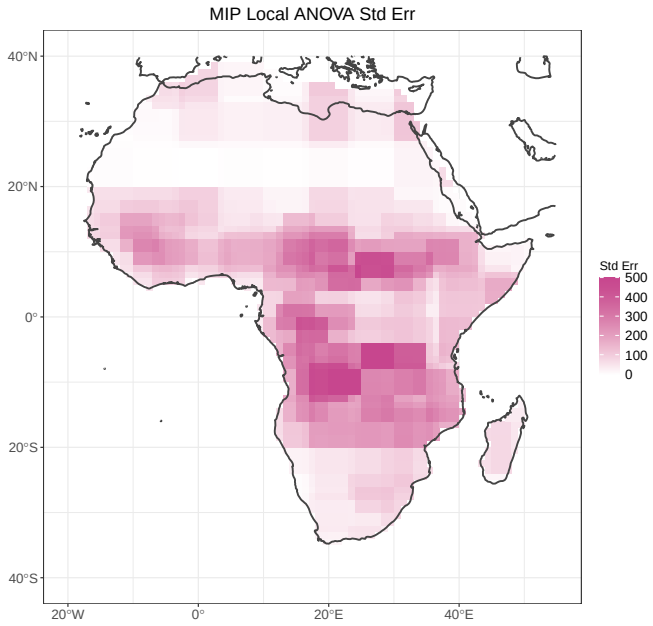

Figure RC.1: Estimated error standard deviation $\hat{\sigma}_\epsilon$ for location-specific ANOVA model fits for the MIP Africa example. Units are gC m^{-2} yr^{-1}

spatio-temporal variability in the flux fields. Even so, processes could exhibit spatially heterogenous variability. We have conducted some exploratory analysis to address the potential for heterogeneous variances, particularly for the ANOVA error term. One tool for this investigation is location-specific ANOVA model fits. These individual results provide an estimate of the error standard deviation $\hat{\sigma}_\epsilon$. These estimates are displayed in Figure RC.1.

This analysis does suggest that the error process variability could change across the region. While the magnitude can apparently be connected to land cover, other factors, including algorithmic choices for the various flux inversion systems likely contribute to this result as well. The prior flux uncertainty could contribute to this behavior as well, but that information is unavailable in the OCO-2 flux MIP output. At this juncture, we do not have a parsimonious and identifiable representation in the model for this heterogeneity. In the revised manuscript, the discussion includes additional remarks on this topic.

- **Also on the topic of oceans, it seems that the fluxes for most of the oceans are zero, or close to zero on the colour scales shown, except for some small areas (e.g., over the Mediterranean or Red Seas in Figure 10). Is this intentional? Is a land-mask being imposed? I recommend discussing this further.**

  Yes, a mask has been imposed, and all of the analyses are for fluxes over land regions only. Each example uses fluxes over specified TransCom regions, as outlined in Table 2. For additional exposition, we have added some discussion of the selected TransCom regions at the end of Section 2.1 and Section 2.2 in the revised manuscript.

1.2 Minor Comments

- **Multiple places: There are multiple references to the supplement of this manuscript. I recommend making the references more specific, such as to particular figures, tables, sections, pages, etc.**

This is an excellent suggestion. For each reference to the supplement in the revised main manuscript, we have included a relevant specific section or subsection.

- **Abstract: I thought this was rather short. I recommend adding a few sentences expanding upon conclusions and implications of this work.**

  The revised manuscript includes an updated abstract. This abstract includes additional discussion of the results, particularly the contrasts among the spatio-temporal covariance parameters for the ANOVA components. The residual term exhibits similar behavior across all of the examples, and this is also mentioned in the revised abstract.

- **L3: "are combined" → "can be combined"**

  This change has been made in the revised manuscript.

- **L4: "impact flux estimates" → "impact such flux estimates"**

  This change has been made in the revised manuscript.

- **L8: I found the term "mode differences" somewhat cryptic. Please review and reformulate.**

  The term has been removed from the updated abstract to provide additional clarity on the results of the CMS-Flux example over Eurasia.

- **L31: The term "data fusion" isn't particularly well defined. I suggest adding some clarifying remarks or using a different term**

  We agree that a definition of data fusion is needed. This has been added at the above location in the revised manuscript. The definition emphasizes addressing spatio-temporal correlation in producing a common estimate at regular resolution/spacing.

- **L37: "ANOVA methodology" → "The ANOVA methodology"**

  This change has been made in the revised manuscript.

- **L41: "multiple component" → "multi-component"**

  This change has been made in the revised manuscript.

- **L47-48: Is the method of Cressie et al. (2022) based on functional ANOVA? If so, I suggest noting this.**

  Cressie, Bertolacci, and Zammit-Mangion (2022) use the classical univariate ANOVA approach for spatially-aggregated flux estimates. In that work, the ANOVA model provides information on partitioning of variability among factors such as the inversion model, but it does not provide any information on the extent of spatial correlation. Therefore we think the current paper provides a meaningful contribution to the flux inversion literature with these additional results using functional ANOVA.

  We appreciate both reviewers raising the importance of framing the functional ANOVA approach and results in the context of the recent contribution from Cressie et al. (2022). In the revised manuscript, we have added a sentence at the above location in the Introduction to highlight our contribution. Further, the connection to Cressie et al. is highlighted again in the first two paragraphs of the Conclusions.

- **Page 2: Reading this far, I was unsure whether the authors were talking about gridded data or not. This became clear later on, but I suggest making a note about this somewhere around this point in the Introduction.**

  This is a good observation. While the statistical model (i.e. Gaussian process) is suitably general for irregular spatial data, the references cited and our work are implemented on gridded datasets. This has been added to the Introduction in the revised manuscript.

- **Page 2: I was left wondering how functional ANOVA handles different spatiotemporal resolutions (or, more generally, different spatiotemporal partitioning). Do all of the data need to be on a common grid?**

  With each of the examples in our study, the spatial resolutions are the same for all factor combinations, but the resolution is different between the MIP examples and the CMS-Flux aggregation example. More generally, functional ANOVA could operate on data with different support/resolution, but an additional change-of-support component would likely be needed in the model.

- **L63-64: References to Sections 2 and 3 are swapped around. Review.**

  The section references have been corrected in the revised manuscript.

- **L69: I suggest noting the units of XCO_2**

  The units (parts per million) have been noted at this spot in the revised manuscript.

- **L70: Do you ave a version number for the ACOS retrieval algorithm?**

  The version number (V9) has been added to the revised manuscript.

- **L73: "within a single polar orbit" - do you mean "within a single model grid-cell"?**

  This phrasing was intended to convey that only retrievals in the same orbit are aggregated together, but the spatial resolution is the more important aspect. This has been changed in the revised manuscript.

- **L75: "a precise" → "a more precise"**

  This change has been made in the revised manuscript.

- **L83: I find URLs within the text, particularly when it is messy such as DOIs, rather unsightly. I suggest creating a bibliography entry for this dataset, and shifting the URL there.**

  We appreciate the suggestion. The dataset has been added as a reference, with the DOI provided in the reference list.

- **L101: "SSDF" - Has this acronym been introduced?**

  We intended to drop the SSDF acronym from the manuscript. The revised manuscript has removed this item.

- **L103: "The variable of interest" → "The variables of interest"**

  This change has been made in the revised manuscript.

- **L122: "the individual retrievals are both variable and moderately correlated in space and time" → "the individual retrievals are both uncertain and the associated errors are moderately correlated in space and time"**

  The suggested modification conveys our intended message effectively and has been incorporated into the revised manuscript.

- **L122: What is meant by "moderately" here?**

  We have added an additional sentence here to mention that empirical analysis of spatial correlation for OCO-2 XCO2 retrievals has found modest (i.e. < 0.4) among neighboring retrievals, plus correlation range estimates on the order of 5-10 km.

- **Figure 1 caption: I don't think the acronym "JJA" has been introduced until now.**

  The acronym is defined in the text, but we have added it to the caption as suggested.

- **L123: "spatially aggregated" - as a gridded or regional average?**

  These are gridded averages. We have noted the approximate along-track extend (75 km) in the revised manuscript.

- **L129: "have similar resolution" → "have similar spatial resolution"**

  This change has been made in the revised manuscript.

- **Figure 2 caption: "over North America." → "over North America, with columns for the different inversions."**

  We have incorporated this additional clarification into the revised figure caption.

- **Table 1: I recommend adding an extra column with a key reference for each retrieval or each retrieval dataset.**

  This additional column has been added in the revised manuscript. In addition, Peiro et al. (2022) and Crowell et al. (2019) contain supporting materials with these references.

- **L166-167: Here we see the issue of months within a season being used for different replicates (see the first of my "Major comments" above).**

  As noted in the response above, the methodology has been modified to incorporate temporal dependence in the error process. We appreciate both reviewers' remarks on this important aspect of the methodology.

- **L170: "are often" - why not "must be"? Is there an alternative?**

  This change has been made in the revised manuscript.

- **L172: "the factor and interaction effects add to zero, e.g." → "the factor effects add to zero, i.e."**

  This change has been made in the revised manuscript.

- **Equations below L172, and L175: Why not put these on the same line? Would they not fit?**

  The equations have been combined into a single line in the revised manuscript.

- **L195: Here the realizations are assumed to be independent (see the first of my "Major comments" above).**

  The revised model with temporal autocorrelation in the error process is discussed at this point in the revised manuscript.

- **L214: θ was not defined, as far as I could see.**

  The revised manuscript includes a definition of θ as the full collection of spatial and temporal covariance parameters at this point in the paper.

- **L215: What do you mean by "diffuse"?**

  We have changed the wording to "high variance" to reflect substantial a priori uncertainty.

- **L223: "for the component and the noise" → "for the components and the noise"**

  We intended this to be "component being updated".

- **Page 11: Here I was hoping to see more about the question of a land mask, or different distributions (e.g., heteroscedastic) for land/ocean areas.**

  Additional discussion of the (land-only) TransCom regions studied has been included in Section 2. Only fluxes over these land masks are analyzed.

- **Page 12: It gets mentioned much later in the Results section (L285) that negative anomalies correspond to carbon uptake by the land, and positive anomalies to the release of carbon into the atmosphere. I think this is an important part of the interpretation of the results shown, and as such suggest introducing this much earlier in the Results section (e.g., within the first few paragraphs, maybe also in the caption for Figure 4).**

  We appreciate this suggestion, as the conventions for fluxes will not be known to all readers. In the revised manuscript, we have decided to introduce this convention for the sign of the fluxes when the datasets are presented in Section 2. We have also added further discussion of the signs of the ANOVA components at the beginning of the results section. The use of the deviations from the prior in the analysis adds some subtlety to the interpretation, and we have emphasized this point at the beginning of the Results section, in addition to the impacts for specific results.

- **L240-245: I would have like to see a bit more interpretation. For example, what are the implications of the estimates for the range parameter being an order of magnitude larger for the mean and year effect than the other terms? What does it mean that the error standard deviation is so much larger than that of the other terms? What can one read from the relatively narrower credible intervals for the error standard deviation than those of the other terms?**

  The revised manuscript includes additional discussion of these results at the beginning of Section 4.1. In particular, we emphasize that the mean and year effects are connected to the behavior of the carbon cycle across the JJA season, and the larger range estimates indicate that the functional ANOVA has estimated spatially coherent perturbations in the carbon cycle over this region. The other terms in the functional ANOVA represent data processing and algorithm characteristics, and they exhibit limited spatial coherenece.

  The width of the posterior credible intervals for the covariance parameters is primarily a function of the number of spatial fields that inform the estimates. This is analogous to degrees of freedom in traditional ANOVA. The availability of the pseudo-replicates yields $n_\alpha\ n_\beta n_\epsilon$ fields to inform the covariance parameters for the error process. On the other hand, $n_\alpha - 1$ fields inform the parameters for α^*, and $n_\beta - 1$ fields inform the parameters for β^*.

- **L254: "The left panel shows $Pr(|\mu(\mathbf{s})| > |\beta^*(\mathbf{s})|)$" - How was that estimated? With reference to the posterior samples?**

  Yes, these probabilities can be estimated from the Markov chain Monte Carlo samples for both processes. We have mentioned this in the text of the revised manuscript.

- **L260: It would be worth noting in the text that the posterior probabilities for the secondary question, $Pr(|\alpha^*(\mathbf{s})| > |\beta^*(\mathbf{s})|)$, are clearly lower than for the primary question, being $Pr(|\mu(\mathbf{s})| > |\beta^*(\mathbf{s})|)$.**

  We have included additional discussion of these combined results in the revised manuscript to address this comment and related comments from the second reviewer.

- **L285: "negative anomalies (increased uptake)" - as noted above, this should be mentioned earlier.**

  In the revised manuscript, we have introduced this convention for the sign of the fluxes when the datasets are presented in Section 2.

- **Figures 7 & 10, bottom panels (data source effect): I would have been interested to see the impact of the locations of the in-situ sites. Consider adding dots for the locations of the CO2 monitoring stations.**

Both reviewers offered a similar suggestion, and we have added in-situ locations to Figures 7, 10, and 11 in the revised manuscript. The implications of in-situ density are discussed for the Africa example, in particular.

- **L295: "The prior fluxes over the continent (not shown)" - Why not put this in the supplement?**

  This is a good suggestion. We have included maps of the prior fluxes for all examples in the revised supplement.

- **Around L293-294: I would suggest noting that for much of the domain, the direction of the flux estimated by the posterior is ambiguous.**

  This note has been added to the results discussion in the revised manuscript. The paragraph in question has been modified further based on the comments from the other reviewer.

- **Last paragraph on Page 15: I would recommend expanding some of the interpretation. On the question of the data sources, adding to the Figure 10 the locations of the monitoring sites could provide some basis for some of the interpretation of the data source effect.**

  The revised manuscript has included some additional interpretation for these results, particularly for the data source effect and the locations of in situ observations. The locations of observation sites have been added to Figure 10.

- **L320: "The functional ANOVA identified local consensus flux anomalies for both continents in the presence of variability across inversion systems and atmospheric CO2 data sources." → "The functional ANOVA identified local consensus in flux anomalies for both continents across different inversion systems and atmospheric CO2 data sources."**

  This change has been made in the revised manuscript.

- **L339: "heterogeneous across space" - I think there's a word missing here. Do you mean heterogeneous standard deviations?**

  Yes, this has been corrected in the revised manuscript.

- **Discussion: I think it would be useful to expand upon the importance of spatial heterogeneity. See the second of my Major comments.**

  This is indeed an important point. In addition to the previous responses, we have added further discussion in the Conclusion to the revised manuscript. This added discussion includes the potential for incorporating land cover information and modeling the error process variance as a function of the prior flux uncertainty. This information is not available in the OCO-2 flux MIP output, but would potentially be provided from flux inversions. A final point is that any modified parameterization would have to be considered carefully in order to maintain identifiability and computational tractability in the inference.

- **References, multiple places: CO2 should be CO_2 and XCO2 should be XCO_2**

  We appreciate the careful reading of the references. The symbols have been modified for consistency in the revised manuscript.

- **References, multiple places: at least two of the author lists are truncated (finishing with "et al.") while others extend to many authors. Please make this consistent, preferably checking the GMD Guide to Authors.**

  The references have been modified to include the complete author lists, as noted in the GMD author guide.

References

Cressie, N., Bertolacci, M., & Zammit-Mangion, A. (2022). From many to one: Consensus inference in a MIP. *Geophys. Res. Lett.*, *49*, e2022GL098277. doi: 10.1029/2022GL098277

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevallier, F., ... Jones, D. B. A. (2019). The 2015-2016 carbon cycle as seen from OCO-2 and the global in situ network. *Atmos. Chem. Phys.*, *19*, 9797–9831. doi: 10.5194/acp-19-9797-201

Peiro, H., Crowell, S., Schuh, A., Baker, D. F., O'Dell, C., Jacobson, A. R., ... Baker, I. (2022). Four years of global carbon cycle observed from OCO-2 version 9 and in situ data, and comparison to OCO-2 v7. *Atmos. Chem. Phys.*, *22*, 1097–1130. doi: 10.5194/acp-22-1097-2022