Reviewer comments in red, responses in black. When necessary for clarity, parts of the previous review are included as indented, italicized text.

The authors believe we had misunderstood many of Reviewer 1's original comments. We hope that the responses below better address the issues this reviewer raised. Notably, 1) we changed from "calibrate" to "bias-correct" as the reviewer suggested, 2) the conclusions were rewritten to address the reviewer's concern about scientific and ACP-specific relevance, and 3) we've provided LoFI NRT comparisons below showing our results hold after 5 years of simulation.

Line 442: "The authors feel ... in our research"
This argument cannot be a justification for publication in ACP.

We have rewritten the conclusions section to give the reader more context for how the results fit into the carbon cycle literature, especially that in ACP, and what the major takeaways for simulations, flux inversion systems, and model-data evaluation are. We hope that the revised text and the responses herein, most importantly the last, address the reviewer's concern about the paper's suitability for publication in ACP.

Line 452: "The authors agree ... technically appropriate"
I don't see why "bias-correcting" would suggest a correction that is constant in time. I would rather think of a calibration as being relatively constant in time (for a good measurement device). A calibration is like a bias correction, but then of a measurement device against a commonly accepted standard. I don't see how it can apply to this case. To which flux standard is being calibrated?

The authors found this along with the previous justification compelling and have changed all instances of "calibrate" to "bias-correct".

Line 459: "The first part ... Line 91"

> *2 The LoFi flux collection: The structure of this section is unclear to me. I had expected three sections, one for the "retrospective mode" on for the "forecasting mode" preceded by everything that is common for these modes. I thought the latter was the baseline, which confusingly enough is not exactly what it turns out to be (see my next point).*
>
> *The first part of the section describes the components of the flux collection, and the second describes what we have to do differently in NRT. The baseline is simply the LoFI flux collection without the empirical land sink, as is noted in the paper on Line 91.*

I was not asking for an explanation, since I have read the paper. The comment was about the clarity of the structure in the eyes of an independent reader. From the answer I conclude that the authors chose to ignore my constructive attempt to improve the manuscript.

The authors agree that the section/subsection structure of Section 2 was somewhat confusing. We have tried to address this by sub-sectioning it as 2.1 "Individual flux components", 2.2 "Anthropogenic short-cycle burning and lateral fluxes", and 2.3 "Modifications needed for forecasting mode". We hope that this division makes it clearer to the reader exactly what each part discusses.

Line 466: "You do understand that correctly ... of this paper"

> *Line 91: If I understand well the baseline still requires the NOAA MBL CO_2 measurements for the ocean flux, which would make it a "retrospective" type analysis. Some explanation is needed of the purpose of the "baseline" other than the notion that it doesn't include the empirical land sink. Initially I was assuming that it would be independent of the NOAA MBL CO_2 measurements, which apparently is not the case.*
>
> *You do understand that correctly. The main purpose of the paper was to evaluate the empirical land sink, so we kept all other products the same in the baseline package. While it would be interesting to test against an alternative baseline that does not use NOAA MBL CO_2 measurements in any way, testing the importance of the MBL CO_2 measurements, especially for constructing ocean fluxes, was not a primary objective of this paper.*

The authors do not get the point that I'm looking for a clearer description here, since I got confused with the current formulations. It needs improvement so that other readers do not need to struggle like me to understand what was done.

We have attempted to clarify this point by explicitly stating that the only difference between the baseline and LoFI fluxes is the empirical land sink.

Line 500: "The purpose ... does well, etc"
The point here was that I was missing an explanation of how the land sink was optimized. This sentence suggested that it was going to be explained in a results section, which is not the right place to explain a method.

This issue has been addressed in the revised text. Now the empirical sink component text in Section 2.1 lays out the exact equations used to compute the sink. We've also edited the text some to help avoid the possible confusion that the reviewer pointed out of the reader thinking the method would be explained later.

Line 495: "It very well could ..."

> *Line 152: ", yet the sink due to the corn and soybean harvest ..." This suggests that the midwest is the sink accounts for the global emission of short cycle fuels.*
>
> *It very well could, but this question is beyond the scope of this paper.*

No, clearly not. There are many other places in the world for which the same holds, so there is no reason to single out the mid-west here.

The authors misunderstood again here the question the reviewer was asking. We have now explicitly stated that our version of CASA-GFED represents the Midwestern US corn and soybean harvest but includes no other harvests. The authors agree that a global representation would be better, and this is the focus of ongoing work at NASA GSFC.

Line 512: “See line 155 ...”
The problem is still there. Section 3.2. is called “the empirical land sink”, whereas it discusses NBE throughout, which is not the empirical land sink.

The authors again misunderstood the point of the reviewer's comment. We have changed the section heading to “Net Biospheric Exchange”. We agree with the reviewer that the evaluation is of the NBE of LoFI, not specifically the empirical sink that we imposed.

Line 539: “The authors feel ... in the paper”
The review is suspicious that this was not shown whereas it could easily have been, because it would raise questions.

An evaluation of LoFI in NRT mode for 2015–2018 against in situ and TCCON data compared to that of modern satellite and in situ inversion systems is available at

https://www.esrl.noaa.gov/gmd/ccgg/OCO2_v9mip/
https://gmao.gsfc.nasa.gov/gmaoftp/sourish/OCO2/MIP/v2/tccon_comparison/plots/summary/

The performance of LoFI in NRT mode relative to the inversions is perhaps better in 2017 and 2018 than it was in 2015 and 2016. While inversion results are not yet available for 2019, LoFI is (a major point of this manuscript). Comparisons for 2019 LoFI NRT against in situ data show comparable performance to previous years. LoFI NRT performance is thus comparable to inversions (even at non-MBL sites) over the 5 years 2015–2019. Furthermore, the submissions in the above intercomparison are the subject of several submitted an in-preparation papers which show, among other things, that LoFI NRT runs have the best representation of the contrast of CO_2 across passing weather fronts as observed in NASA’s ACT-America suborbital campaign. While these results are compelling, they are not included in the manuscript because they are the subject of an intercomparison-wide paper in preparation analogous to Crowell et al. (2019; doi:10.5194/acp-19-9797-2019) for a previous OCO-2 retrieval version.

One point that is important to note is that these runs are only forecasts of carbon fluxes. All runs use reanalysis meteorology from MERRA-2 which is available roughly 1 month after the current date. If we were to have forecasted the meteorology in addition to carbon fluxes, we would expect the skill to degrade within a few months and last about two years or less as demonstrated by Ilyina et al. (2020; doi:10.1029/2020GL090695). This point has also been included in Section 3.4.