Biogeosciences Discuss., referee comment RC1
https://doi.org/10.5194/bg-2022-160-RC1, 2022

Comment on bg-2022-160

Anonymous Referee #1

Referee comment on "Resolving scale-variance in the carbon dynamics of fragmented, mixed-use landscapes estimated using Model-Data Fusion" by David T. Milodowski et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2022-160-RC1, 2022

General comments:

This paper demonstrates the impact of using both finer-scale and categorically refined representations of land surface heterogeneity on modeled carbon stocks and fluxes. The authors present a case study over a region with four dominant land use types to demonstrate that modeling the ecosystem response of each land use type separately, and aggregating the results, does not always yield the same result as modeling the aggregate ecosystem response of the region. The authors document differences in simulated carbon stocks and fluxes, and derived parameters characterizing the ecosystem function, among the different approaches and resolutions tested.

The hypotheses are interesting and well-explored by the experiments chosen, and the results are important. By documenting the sensitivity of the data assimilation framework to the spatial scale and categorization of the data inputs, this work highlights ecological assumptions embedded in standard usage of this and similar models that may undermine their ability to investigate questions of ecological function and future response. Raising this issue is a useful contribution.

However, some of the framing and concluding statements in this paper assert improvements in ecological fidelity or simulated carbon fluxes due to stratification without validating this with any outside data. It would be great if possible to include some validation, such as comparison with outside data to validate derived parameters (e.g., residence times) or carbon/water flux data from flux towers. If this isn’t possible, I recommend adopting new language in your framing and conclusions to focus on the sensitivity you demonstrate and make less claim to ecological fidelity or improved representations of carbon fluxes. Section 4.3 is informative and some of the context outlined there could be brought out in the framing and goals at the outset.

Specific comments are below:

Comment 1: Section 2.3-2.4: It would be helpful to see a paragraph at the end of the methods discussing the various spatial scales at play, and how these are integrated into the model pixel in each case. When working with a 0.05 degree (~5km) model resolution but imposing constraints on biomass at 100m, on a soil type at 250m, an LAI at 300m, and a timber harvest at 30m, how are these aggregated across the pixel?

Comment 2: As a follow on from comment 1, the description of study area emphasizes gradients in temperature and precipitation over topographic features within the 3x3 grid (Lines 128-131), which justifies testing surface resolutions down to 0.05 degree, but then the model runs use 0.5x0.5 forcing data in each case. How do the authors expect this to relate to the amount of scale-dependent variation seen across the model runs?

Comment 3: Section 3.1 The calibration metric, RMSE/sigma, could be further explained. It sounds as though smaller values are desirable here, but if this is a comparison to inherent observational uncertainty, I don't immediately see why <1 is a good thing. Please make this a bit clearer.

Comment 4: Table 1, Table A2: Additionally, the values of the calibration metric do not proceed monotonically with the shift in resolution. It would be helpful if the authors could explain (or speculate) why this is, especially in the context of the stated goal of improving ecosystem representation by going to smaller scales. A response to this could connect to a response to comments 1 & 2— how does the scale of the input data impact how well things are lining up in the model (applying the right processes to the right initial conditions) at different resolutions?

Comment 5: Line 275 and Figure A2: Please explain why the coniferous woodland has a strong seasonal cycle of LAI which reaches zero in the winter. The black dots in Figure A2 top left panel suggest that this oscillation is present in the earth observation data, but Scotland is not known for its deciduous conifers. Does snow blanketing the tree canopies, masking out the greenness or making the canopies indistinguishable from the ground, cause this seasonality in the Copernicus LAI product, or is this considered ecologically realistic for your region? If it is unrealistic, does this matter to the resulting biomass trends— for instance, did the authors test a different (presumably more realistic) oscillation bottoming out at ~3?

Comment 6: Section 3.3: Regarding the differential response to disturbance flux, it would be helpful if the authors emphasize earlier on that this arises from a mismatch in applying the disturbance to the correct land cover type when using the aggregated pixels. I see the authors do come to this in lines 358-361 but would appreciate it earlier. Section 2.3.5 could be a good place to explain how the authors imposed the disturbance flux in each case, so it is ultra-clear why this difference in how the disturbance is allocated to each land use type arises between the two cases.

Comment 7: Section 3.3: Line 319: "it is evident that stratification leads to preservation of ecological information across resolutions" and similar statements throughout; suggest to qualify these statements, e.g., "as encoded in the observations available to the model". Getting back to the conifers acting like deciduous trees— it is important to tread carefully with caveats that the observations themselves come with many assumptions, and may not always represent ecological fidelity. The authors make this caveat in section 4.3 line 395-400, but it would be helpful to keep it at the forefront throughout.

Comment 8: Section 4.3 line 395-320: Great points. It would help satisfy reader curiosity if the authors could delve a bit into these deficiencies (Zhao et al 2020, Heiskanen et al 2012) as relevant to the datasets they are using, and discuss how the deficiencies might impact their results. This comment has substantial overlap with comment 5.

Comment 9: Line 406-415: This is a very strong point. It would be great to go further and see the authors chart out a bit what is needed to actually do these improvements in process representation— what is the to-do list? How will improvements be verified?

And a few minor **technical comments** below:

Comment 10: Table 1 last CWood stratified row, shifted numbers

Comment 11: Figure 6: What does the color of dots (blue vs green) in the right panels mean? A legend would help

Comment 12: Figure A2-A3: A legend would help here also.