



EGUsphere, referee comment RC1
<https://doi.org/10.5194/egusphere-2022-641-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on egusphere-2022-641

Anonymous Referee #1

Referee comment on "Towards an ensemble-based evaluation of land surface models in light of uncertain forcings and observations" by Vivek K. Arora et al., EGU sphere, <https://doi.org/10.5194/egusphere-2022-641-RC1>, 2022

Review of egusphere-2022-641:

Towards an ensemble-based evaluation of land surface models in light of uncertain forcings and observations

Summary

The authors present an evaluation of their updated land model, CLASSIC. The primary update from a previous model is a new nitrogen cycle, although new land cover reference data have been implemented also. Notably, they evaluate the model using two different meteorological driving data sets, and also comparing new vs old land cover data and with/without the new nitrogen cycle. Their evaluation system effectively compares model scores to benchmark scores that are based on observational uncertainty. They conclude that their new model is reasonable compared to both observation and other models, and also that present-day land atmosphere co₂ flux is independent from the initial land carbon state, with respect to the variations included in this experiment.

Overall review

I appreciate the authors' more comprehensive approach to evaluating their updated land model. The evaluation system provides clear results. However, in the end it isn't clear that the model is better, but it appears that the GSWP3 forcing gives better results than the CRU forcing. I also think that the conclusions regarding the independence of flux from initial land state are overstated, mainly because this is a highly constrained case where the land start and end points are shifted by a similar amount and the land change trajectory is nearly identical (but shifted) between the two cases. I recommend the following main revisions (see below for additional details):

1) The different land cover cases need to be redefined. They are not different reconstructions. They just reflect an update in the present-day reference land cover data that are used to anchor the model's land trajectory. While this is a reasonable update, it does not represent the uncertainty of land use/cover change.

2) Qualify your conclusions regarding the robustness of model fluxes under different initial carbon states. This is a very specific, highly constrained comparison and there are many factors and uncertainties, particularly in the land space, that are not considered here but have substantial impacts on carbon flux and storage estimates.

3) Complete your background and comparisons with literature on land data uncertainties. Some suggestions are below, but there is more out there showing the complexity of this problem.

4) Swap the more useful appendix figures for the unreadable paper figures. Try to make

the figures more readable.

Specific comments/suggestions

Abstract

line 6:

I would not call these two land cover sets different historical reconstructions. You just replaced your current day reference land cover with newer data. There are many other factors that affect the reconstruction, most notably the land use data and the assumptions used to apply land use to land cover.

line 12:

awkward sentence transition. probably do not need the beginning of this sentence; just start with "Simulated area burned..."

Introduction

lines 90-94:

there are other studies on this topic. the most relevant one is probably this one because it addresses uncertainty in land cover in conjunction with the effects of co2, nitrogen deposition, and climate:

A.V. Di Vittorio, J. Mao, X. Shi, L. Chini, G. Hurtt, and W.D. Collins, "Quantifying the effects of historical land cover uncertainty on global carbon and climate estimates", *Geophysical Research Letters*. doi: 10.1002/2017GL075124.

This one looks at land change emissions across several land cover representations:

Peng, S., P. Ciais, F. Maignan, W. Li,

J. Chang, T. Wang, and C. Yue (2017),

Sensitivity of land use change emission

estimates to historical land use and land

cover mapping, *Global Biogeochem.*

Cycles, 31, 626–643, doi:10.1002/

2015GB005360.

but this one is also relevant:

A.V. Di Vittorio, X. Shi, B. Bond-Lamberty, K. Calvin, A. Jones, 2020, "Initial land use/cover distribution substantially affects global carbon and local temperature projections in the integrated Earth system model", *Global Biogeochemical Cycles*. doi: 10.1029/2019GB00683.

CLASSIC modeling framework

lines 178-180:

Some folks may disagree here. Different types of trees have different leaf/canopy shapes, orientations, and colors that may affect interception and also radiative processes.

Driving data

line 230:

This may be true in some cases, but the later step of creating the reconstruction by applying the land use trajectory to this static cover map can generate greater uncertainty, not to mention the additional uncertainty in the land use data. See the papers above. See comment below. And also these:

Di Vittorio, A.V., L.P. Chini, B. Bond-Lamberty, J. Mao, X. Shi, J. Truesdale, A. Craig, K. Calvin, A. Jones, W.D. Collins, J. Edmonds, G.C. Hurtt, P. Thornton, and A. Thomson (2014). From land use to land cover: restoring the afforestation signal in a coupled integrated assessment - earth system model and the implications for CMIP5 RCP simulations, *Biogeosciences*, 11:6435-6450, 2014, doi: 10.5194/bg-11-6435-2014.

Meiyappan, P., and A. K. Jain (2012), Three distinct global estimates of historical land-cover change and land-use conversions for over

200 years, *Front. Earth Sci.*, 6(2), 122–139, doi:10.1007/s11707-012-0314-2.

lines 266-286:

Figure 1 indicates that changing your reference cover map does not generate the greatest uncertainty. The range of vegetation across the other models is much greater than the difference you show for your data. What is the nominal year for your data sets? What about for the other models? What is actually driving the variability in these data across trends? Are they all using different reference data? Or are other factors contributing?

lines 358-387:

This section is unclear, particularly with respect to how the benchmark scores are calculated (the ones comparing the obs). You show the benchmark scores in figure 10, but it isn't clear how these are calculated. I do like this benchmarking system, though.

line 393:

figures a2-a16 are much more useful than figures 3-9, which are unreadable.

Results

lines 558-559:

It is unclear how the benchmark scores are determined.

Conclusion

lines 649-651:

The key word here is "present-day flux." The cumulative emissions over time are dependent on the land cover change trajectory, which you do not alter in these scenarios. This is also why your previous statement regarding model response being independent of initial land state makes sense here; you do not have a different transient land path that would change the outcome. Your two land covers are both estimates of "present-day" cover, and as such are not that different from each other. And using the same land cover backcasting your initial state changes by a similar amount. So the flux is also constrained by similarly adjusted endpoints.

Note that there are grammatical typos throughout.

Figures and Tables

Figures 3-5:

I cannot determine which simulations are where. A couple of colors are clear, but the groups of lines are muddled together and I cannot tell which simulations are in which group. If you colored them by output group it would be easier to tell which sims have similar results. Using a temporal average may also help (without the annual values shown, which make it messy).

Figures 6-9:

These are difficult to read. Since output groupings are less apparent here, I suggest selecting colors that reflect the experiment groupings. Temporal averages may also help here (without the annual values shown).

Figures A2-A16:

make sure the axis scales match across all panels in each figure.