Comment on esd-2020-92

Anonymous Referee #2

Referee comment on "Comparison of uncertainties in land-use change fluxes from bookkeeping model parameterisation" by Ana Bastos et al., Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2020-92-RC2, 2021

Review of esd-2020-92:

Comparison of uncertainties in land-use change fluxes from bookkeeping model parameterization

Summary

The authors present an analysis of differences between two bookkeeping models of land change emissions. They replace parameters in one model with those from the other model in a factorial design to isolate parametric sources of the differences. They find that the drivers of differences are, in descending order, harvest allocation, carbon density, net vs gross transitions, and decay time scales. They conclude that uncertainties in these drivers are too high to determine which model has the best parameters, and that further data development is critical for assessing land change emissions.

Overall impression

This is a very interesting paper that sheds some light on sources of uncertainty in land change emission estimates. For the most part it is clear and straightforward. However, I find it incomplete in three respects: lack of model detail and discussion regarding the processes studied, omission of the effect of initial vegetation state, and it leaves me with the question: why are BLUE emissions lower than HN2017 when using HN2017 parameters? Addressing these concerns will strengthen the paper. I flesh them out here, and there is more detail below.

1) There needs to be more description of and subsequent discussion of the particular sources of uncertainty and their effects, particularly for the harvest allocation parameters and how they are implemented, which dominate the sensitivity. The gross fluxes also need to be explained because their results don't seem to match well with intuitive understanding of the processes controlled by the parameter uncertainty groups.

2) The initial vegetation distribution plays a major role in subsequent carbon emissions, and interacts directly with carbon densities and land change to determine the emissions and carbon state of the system at any point in time. There are three distinct initial states in this study: BLUE start (850), HN2017 start (1700), and the accounting start (1850). Each model has its own state at these points, and the differences in emission estimates cannot be separated from these different states. This is partly because a mapping between the vegetation distribution and the prescribed transitions has to be made (and is different for each model), and potentially because the prescribed transitions may not be fully carried out in the models due to limited availability of particular land types (especially for BLUE). I am not sure whether HN2017 takes the vegetation distribution into consideration, or just assumes that prescribed changes occur in full. Consider taking a look at this paper:

Di Vittorio, A. V., Shi, X.,Bondâ□□Lamberty, B., Calvin, K., & Jones,A. (2020). Initial land use/cover distribution substantially affects global carbon and local temperature projections in the integrated earth system model. Global Biogeochemical Cycles, 34, e2019GB006383. https://doi.org/10.1029/2019GB00638

This can also affect carbon densities if at any point they are averaged across mature and recovering land types (because of different area weightings). In any case, this is a factor

in understanding the questions in the following concern.

3) Blue estimates much higher emissions than HN2017, but when BLUE uses HN2017 parameters BLUE estimates lower emissions. This clearly shows that the parameters isolated here are not the only sources of difference. It is implied in the paper that the different LUC forcing data account for this, but it is never quantified or shown. Other statements indicate that this different forcing is a cause of the high emissions, but the SHNfull simulation suggests that the LUH net forcing results in less emissions, given the same parameters. Is this because the net changes are less in the LUH forcing than in HN2017? What other model differences contribute to this?

Specific comments/suggestions:

Abstract

Introduction

lines 57-59:

Is this a normative statement of how it should be done? Or is this a condition based on how these estimates were made? And were the DGVMs designed to capture such BK uncertainties, or is this just a hope that their differences would reflect those of the BK modes?

Data and Methods

lines 92-110 (also table 1):

What is the baseline land cover for each? It is likely different for each, and a major contributor to differences in outputs, but you do not describe it in the text or include it in the table. You also discuss the PFTs in the next paragraph, and that their distributions may differ, but do not state where the PFT distributions come from or how they differ between the models. You also indicate a potential vegetation map in figure 1.

lines 161-176:

It would be helpful to clarify in each of these descriptions that the 1700-net setup is also used.

lines 168-176:

More description is needed for these two. Are the slash fractions associated with other pools, or are they their own pools? What is the time scale of HN2017 slash decay?

These allocations seem to be generate the greatest differences, so how they operate in the models should be explained.

Are these decay times for all (and only) the pools associated with the alloc experiment

(the pools listed in table A3)? It seems there may be other pool affected based on figure 5, otherwise, how could they make such a difference for crop-pasture transitions?

lines 198-199:

I don't understand what this is.

Results

line 206 and figure 2:

It isn't clear that you did an HN2017 run. You should clarify this in the methods.

lines 227-238:

What about the differences due to the Alloc simulation? What about this allocation causes the largest difference of the three parameter sets? Which pools

lines 249-254:

this is difficult to follow. maybe try a descriptive sentence for each region.

Also, RUS contributes to this, and NSA clearly contributes to this, even though it may have a smaller magnitude than the others.

line 264:

agreement with HN2017, correct?

lines 267 and 271:

this is not an intuitive way to assess this, so it is confusing. it also isn't clear how this is calculated, which contributes to my confusion. But I think it is easier to understand if you showed the rmsd to hn2017 for each simulation, in which case it would be lowest for alloc and highest for t in fig 4b. The reader then interprets this as lower rmsd is closer to the reference. As it is now, the greatest decrease in difference for alloc with respect to the net simulation is a bit convoluted. Likewise for the full simulations. You may want to add the standard Sbl sim values to the top part of 4b if you change this.

lines 299-319:

this does not seem to be consistent with the previous results, except for compensation by crop-pasture transitions, which for some reason have strong responses to these parameters. I think you need to describe how these gross fluxes are grouped because they don't seem to align with the parameter groups. For example, what to harvest allocations have to do with abandonment and crop-pasture transitions? I would expect the differences for alloc to be in the wood harvest flux group. Likewise time decay constants: aren't these related to harvest allocation pools? How do these affect crop-pasture transitions?

And I think that showing percent difference in figure 5 can be misleading, as the magnitude of each flux group can be dramatically different, for example a 60% difference in crop-pasture fluxes may be similar in magnitude to a 10% difference in clearing emissions.

Discussion

lines 352-354:

The differences in luc forcing should be quantified in this study. On the one hand, it contributes to higher emissions in blue, but if blue is parameterized like hn2017 then blue has lower emissions. So how does the same forcing difference drive these two opposite results? It seems that the LUH data may have less net transitions than hn2017, which would mean that this luc forcing is not a driver of higher emission estimates (although gross transitions do increase emissions, as expected).