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

Ana Bastos, Kerstin Hartung, Tobias B. Nützel, Julia E.M.S. Nabel, Richard A. Houghton, Julia Pongratz

Response to reviewers

Reviewer #1

This paper separates the effects of land-use change, carbon density, allocation of carbon pool through an interesting controlled simulation of two bookkeeping models. This study is useful to clarify the sources of uncertainty in the calculation of CO2 emission from land-use change, which is an important topic and addressed nicely by this paper. However, some methodology and parameters are not very clear and need improvements. Please see specific illustrations as below:

1. The cooperation of SBLand SBL-Net shows the impact of land-use transitions. However, how about the difference of spatial distribution of land-use of LUH2v2.1 and FAO. If SBL-Net is 13% lower than SBL, in what probability is it affected by the difference of land-use distribution of LUH and FAO?

AR: Indeed, the spatial patterns will differ between the two models. HN2017 is based on FAO data, which have only national information. LUH is partly based on FAO (for cropland and pasture), but is further combined with remote-sensing (recent decades) and national statistics (before 1960). To isolate the effect of net vs gross we stay consistently in one framework (LUH). In FAO, most of the net vs gross could not even be captured, because shifting cultivation and other sub-national gross changes are not considered.

There are further differences arising from the potential vegetation map used, and different PFT classes. In BLUE, LULCC transitions from LUH2 are overlain by a potential vegetation map following (Pongratz et al., 2008), while H&N estimates initial areas of crop, grassland, forest and other vegetation and then propagates transitions. Therefore, BLUE and H&N areas of crop, pasture and forests at any given time-step will not necessarily be the same. In the revision of the MS we include a comparison of global areas of these four vegetation types in 1700.

2. Figure 2 shows the global flux from land-use, what factors contributes to the differences of S_{HNFull} and HN2017?

AR: BLUE and HN2017 differ in several aspects beyond the model parameters, as summarized in Table 1. While here we test differences due to parameter choice, these are then amplified/dampened by interactions with model structure (e.g. linear vs. exponential response curves) and LUC transitions (see our reply to Reviweer #2, point 3 below). We agree that this needs to be clarified in the manuscript and will revise the text accordingly.

3. This paper only analyzes the contributions of land use transitions, carbon density and allocation. Other factors such as climatology also affect the FLUC. How about the contribution of other factors in addition to those three parameters? I suppose these two BK model use the same climate forcing data?

AR: Bookkeeping models do not include climate or other environmental changes to estimate FLUC. This is a known feature of bookkeeping models (Pongratz et al., 2014), and has been addressed in detail in the companion paper by (Obermeier et al., 2021).

4. For Figure 5, it is a bit confused tome that crop-pasture transitions result in 60% FLUC difference in Sahara desert. Crop-pasture transitions is common in Sahel, but Sahara desert should rarely have crop-pasture transitions?

AR: This is an artifact that results from showing the results as %, rather than in absolute terms. Small differences can lead to unrealistically high values in relative terms if the fluxes in the denominator are very small. We will update the figure and mask out areas where FLUC are very small, thereby leading to extremely large differences that have little physical meaning.

5. How about the uncertainty of carbon allocation parameters used by the two models?

AR: This is an important uncertainty that we address in lines 400-409. However, contrary to C densities, at the moment no global dataset of allocation parameters exists that could be included in our simulations. We will stress this in the discussion.

Reviewer #2

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.

AR: We thank the reviewer for the encouraging evaluation of our study and for the constructive comments. We address these comments in our replies 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.

AR: We recognize that the model and parameter description were not fully detailed. We chose to reference the recent literature about other sources of uncertainty to keep the manuscript (already long) within a reasonable length to improve readability. However, we realize now that this comes at the cost of readers having to refer to the original manuscripts. We will add more detailed descriptions in a new section in Annex.

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.,Bond-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.

AR: thank you for the reference, which is definitely relevant for our study. Indeed, BLUE and H&N areas of crop, pasture and forests at any given time-step will not necessarily be the same, since BLUE LULCC transitions are overlain on a potential vegetation map following (Pongratz et al., 2008), while H&N fixes initial (1700) areas of crop, grassland, forest and other vegetation (see description of methods in Houghton & Nassikas 2017). In the revision of the MS we include a comparison of global areas of these four vegetation types in 1700. We will also include a deeper discussion about the relevance of the background potential vegetation, including the reference proposed.

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?

AR: Indeed, parameters alone explain only a fraction of this difference, as summarized in Table 1. The differences between the two bookkeeping models can be due to:

(i) differences in forcing data. This has been addressed by (Gasser et al., 2020), (Grassi et al., 2018) and, to some extent, also by (Bastos et al., 2020). Recently, (Hartung et al., 2021) estimated the uncertainty due to different LUC scenarios, including differences in starting years. They showed that differences in the forcing have larger influence than the starting years. (Bastos et al., 2020) have also shown that differences between LUH2 and the FAO forcing contributed strongly to uncertainties in regional carbon budgets. Given that these differences have been previously reported and thoroughly investigated, and that their inclusion would require a very large number of factorial simulations to track interactions between parameters and forcings, we did not include them in this study. It is not likely that the lower values result from lower net changes in LUH. Gasser et al. (2020) has compared FLUC resulting from different forcings, and LUH2 led to higher emissions than FAO/FRA globally (regionally there are important differences between forcings).

(ii) differences in model structure. This is described in Section 2.1 and summarized in Table 1. Calculating the contribution of these differences in model structure to the uncertainties in the resulting fluxes would require changing the model's structure. Given that none of the two models was built in a way that facilitates updating or exchanging particular modules, this is not realistically feasible.

(iii) differences in model parameterization. Both models used observation-based parameters, but these have high uncertainties. To the best of our knowledge, no study has estimated how uncertainty in parameters propagates to FLUC, which is the reason why we focused on this aspect. However, as discussed in the manuscript, explaining difference in FLUC due to changes in parameters is not always easy to track, because errors can be amplified/dampened by interactions with land-cover distribution, differences in transition types, etc. However, while preparing our replies to the reviewers, we noticed that there was an error our conversion of the parameters from HN2017 to BLUE for the allocation of the harvest pool fractions which resulted in unrealistic values (0.00 allocation to the harvest 1-yr product pool, see Table 3). This is very likely to explain the much lower FLUC from BLUE when using these parameters. We corrected these values and are now running new simulations to update the manuscript.

(iv) differences in terminology and definitions and factors included. Here we consider the effect of net vs. gross transitions. Other sources of uncertainty are important when different types of models/approaches are used, but not for comparing these two BK models (Pongratz et al., 2014; Obermeier et al., 2021).

We do not aim to do a full accounting of all the uncertainty terms, but rather to show how resulting global and regional FLUC may depend on parameters that are, themselves, uncertain. An important point to note is that some of these parameters can now, in principle, be constrained by earth-observations. Therefore, our study provides a good motivation to improve BK model parameterization based on remote-sensing data.

We will improve the discussion about these aspects in the revised version of the manuscript.

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?

AR: This is how these estimates were made. We will clarify in the revision.

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.

AR: As mentioned in our reply above, the potential vegetation map in BLUE comes from (Pongratz et al., 2008). We will add one table comparing the distribution of vegetation types in 1700 for both models.

lines 161-176: It would be helpful to clarify in each of these descriptions that the 1700-net setup is also used.

AR: We will add this clarification.

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?

AR: Indeed, these were specified in the Table A3, but not mentioned in the text. We will correct this. As for the pools, yes, these are the only pools affected. The fact that we show results in relative terms, rather than absolute may partly explain why the differences seem so large, in those regions with small FLUC, e.g., the high values in % over the Sahara desert. These high values have, however, little meaning, and therefore we will mask out from the maps regions with very small FLUC (Reply to Reviewer #1, point 4). However, difference in allocation rules and decay times can propagate in time through the slow carbon pools (soils and recovery pools). If the decay constant of wood harvest changes, the amount of carbon still present after a transition will be different and this will impact the emissions from clearing in turn.

lines 198-199: I don't understand what this is.

AR: It is the equivalent of a root mean square error, but we do not refer to it as error because there is no "true value". We will add the corresponding equation.

line 206 and figure 2: It isn't clear that you did an HN2017 run. You should clarify this in the methods.

AR: thank you for pointing this out. We used the run from (Houghton and Nassikas, 2017). We will clarify this.

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

AR: We performed one additional simulation where we had kept the slash pools from BLUE and changed only the allocation pools. Slash fraction differences (more slash in BLUE lines 350-351) did no lead to significant differences in the resulting FLUC, and therefore we included them in the Alloc simulation. Separating each of the other allocation pools individually is not straightforward since the three pools need to add up to one. Given these challenges, and that little insight is likely to be gained from such simulations (given the interactions with other terms, as discussed above), we prefer to maintain the current simulation protocol.

However, while addressing the replies to R#2, we noticed that there was an error our conversion of the parameters from HN2017 to BLUE for the allocation of the harvest pool fractions which resulted in unrealistic values (0.00 allocation to the harvest 1-yr product pool, see Table 3). We will correct this in the revised version of the manuscript, including all the associated figures and discussion.

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.

AR: Thanks for the suggestion, we will revise the text accordingly.

line 264: agreement with HN2017, correct?

AR: Yes, it will be revised.

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.

AR: We will revise accordingly.

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.

AR: We agree with the reviewer. One reason why these differences appear so large is that the percentages can be unrealistically high if fluxes are very small. This explains why in %, some differences appear much larger than they actually are. We will correct this, see Reply to Reviewer #1, point 4.

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).

AR: These differences have been quantified elsewhere (see our reply above), and we would therefore not make them the focus of this study. However, we agree that this needs to be discussed in more detail in this section, so we will revise the manuscript accordingly.

References

Bastos, A., O'Sullivan, M., Ciais, P., Makowski, D., Sitch, S., Friedlingstein, P., Chevallier, F., Rödenbeck, C., Pongratz, J., and Luijkx, I.: Sources of uncertainty in regional and global terrestrial CO2-exchange estimates, Global Biogeochemical Cycles, 2020.

Gasser, T., Crepin, L., Quilcaille, Y., Houghton, R. A., Ciais, P., and Obersteiner, M.: Historical CO 2 emissions from land use and land cover change and their uncertainty, Biogeosciences, 17, 4075–4101, 2020.

Grassi, G., House, J., Kurz, W. A., Cescatti, A., Houghton, R. A., Peters, G. P., Sanz, M. J., Viñas, R. A., Alkama, R., Arneth, A., Bondeau, A., Dentener, F., Fader, M., Federici, S., Friedlingstein, P., Jain, A. K., Kato, E., Koven, C. D., Lee, D., Nabel, J. E. M. S., Nassikas, A. A., Perugini, L., Rossi, S., Sitch, S., Viovy, N., Wiltshire, A., and Zaehle, S.: Reconciling global-model estimates and country reporting of anthropogenic forest CO2 sinks, 8, 914–920, https://doi.org/10.1038/s41558-018-0283-x, 2018.

Hartung, K., Bastos, A., Chini, L., Ganzenmüller, R., Havermann, F., Hurtt, G. C., Loughran, T., Nabel, J. E. M. S., Nützel, T., Obermeier, W. A., and Pongratz, J.: Net land-use change carbon flux estimates and sensitivities – An assessment with a bookkeeping model based on CMIP6 forcing, 1–34, https://doi.org/10.5194/esd-2020-94, 2021.

Houghton, R. and Nassikas, A. A.: Global and regional fluxes of carbon from land use and land cover change 1850–2015, 31, 456–472, 2017.

Obermeier, W. A., Nabel, J. E. M. S., Loughran, T., Hartung, K., Bastos, A., Havermann, F., Anthoni, P., Arneth, A., Goll, D. S., Lienert, S., Lombardozzi, D., Luyssaert, S., McGuire, P. C., Melton, J. R., Poulter, B., Sitch, S., Sullivan, M. O., Tian, H., Walker, A. P., Wiltshire, A. J., Zaehle, S., and Pongratz, J.: Modelled land use and land cover change emissions – A spatio-temporal comparison of different approaches, 1–43, https://doi.org/10.5194/esd-2020-93, 2021.

Pongratz, J., Reick, C., Raddatz, T., and Claussen, M.: A reconstruction of global agricultural areas and land cover for the last millennium, 22, https://doi.org/10.1029/2007GB003153, 2008.

Pongratz, J., Reick, C. H., Houghton, R., and House, J.: Terminology as a key uncertainty in net land use and land cover change carbon flux estimates, 5, 177–195, 2014.