

MS Title: Effect of land use legacy on the future carbon sink 1 for the conterminous U.S.

Author: Benjamin S.Felzer

General comment

The manuscript entitled "Effect of land use legacy on the future carbon sink for the conterminous U.S." by Benjamin S.Felzer assesses the effects of land use and land cover change (LULCC) from 1750 to 2014 using the Terrestrial Ecosystems Model (TEM) for the conterminous U.S. Comparing model outputs from four experiments, the author finds that LULCC legacy has a considerable impact on carbon pools and fluxes. While the topic is of general interest, I don't follow the experimental design of the study. My primary concern is that the author combines meteorological forcing data from one climate model (MPI-ESM-P 1750-1900) with quasi-observed data (CRU, 1901-2014) and climate projections from another model (CCSM4, 2015-2099) in one continuous simulation. Each data set comes with their own sets of biases. This lack of consistency will manifest as a forcing, causing changes in carbon fluxes. So the final signal is a combination of changes in environmental conditions, such as climate change, as well as the transition between inconsistent data sets. Also, one of the experiments (CONDENSED) has been re-equilibrated to the environmental conditions of 2015-2045. It remains unclear to me why the author chooses to equilibrate the model with projected climate conditions when trying to isolate the role of LULCC. I recommend that the author either adjusts the methods, or provides additional explanations that justify his approach. I recommend that the manuscript may be considered for publication in Biogeosciences after major revision.

Detailed Comments

L19 It is common practice to account for LULCC that started during the pre-industrial period. For instance, the TRENDY model ensemble that informs the annual publication of the global carbon budget accounts for LULCC starting in the year 1701.

(see <https://blogs.exeter.ac.uk/trendy/protocol/>)

L41 To my understanding you don't compare model output against observation-based reference data. I would therefore not write that "carbon stocks are overestimated". Instead, I would either describe how carbon stocks differ among experiments or expand the analysis by comparing results against observations.

L55 Replace "address" with "addresses".

L80 Would a carbon sink related to regrowth not be larger if disturbance rates reduce, rather than "continue"?

L80 Spell out the FIA acronym.

L149 This section describes the different experiments (historical, restart, condensed, and tem-restart) with respect to their initial values and whether they are based on the full or condensed version of the cohort. Please add how you treat LULCC, atmospheric CO₂ concentrations,

nitrogen deposition and nitrogen fertilization when describing each experiment, and include this information in the table.

L156 Please motivate why you condense the full cohorts to 1 cohort/PFT.

L166 You write that "the difference between the RESTART and CONDENSED runs shows the effect of including land legacy on future carbon dynamics". The difference between both runs is that the CONDENSED run re-equilibrates using climate data from 2016-2045 (L210). If the CONDENSED run has been re-equilibrated, then it is also in equilibrium with respect to the meteorological forcing, CO₂, and N deposition + fertilization that correspond to 2016-2045. How do you separate the impact of land legacy from these other factors then? Please clarify or adjust your method.

L183 You Combine climate model data from one model, (MPI-ESM-P 1750-1900) with quasi-observed data (CRU, 1901-2014) and climate projections from another model (CCSM4, 2015-2099) in one continuous simulation. The more conventional approach is to conduct simulations that are either based on quasi-observed data or on data from one climate model. As for the quasi-observed climate data you could have used an early chunk of the historical data (e.g. 1901-1920) and spun up the model by iterating this climate data for whatever period it takes to equilibrate your model. This would have also freed you from the need of bias correction and downscaling MPI-ESM-P. The problem with your approach is that you combine data that come with their own set of biases and thereby mix the forcing from environmental factors with differences between these data sets. Please provide an explanation that justifies your experimental setup or adjust your method.

L197 It would be more convincing if you had used radiation rather than cloud data for the historical period as well. The change in your method from one period to the next may create an unnecessary artificial forcing, which then mixes with the impact of climate change. Please justify your approach or adjust your method. t

L210 You write that the CONDENSED run is first equilibrated based on repeated use of the 2016-2045 climate. It is not clear to me why you use projected future climate conditions to equilibrate your model. Please explain or adjust your method.

L211 Please define NCE before using the acronym.

L218 Please mention in the text what experiments you are referring to.

L238 You write that "reinitializing each grid is based on the assumption of NEP as close to zero". Should it not be Net Biome Productivity (NBP) rather than NEP that should be close to zero, as NBP also includes fluxes associated with disturbances, such as wildfires?

L275 Please explain why you expect that more mature forested in CONDENSED would have lower NEP. Also, please mention what period you are referring to. Finally, replace "mature forested" with "mature forests", if this is what you mean.

L278 The CONDENSED run has been re-equilibrated to the environmental conditions of 2015-2045. Is that not the main reason why vegetation carbon is 16% larger compared to the RESTART simulation? Also, I don't recommend using the term bias here if you are not comparing against observation-based reference data (here and elsewhere).

L280 Explain exactly what you mean by "fixes most of the problem".

L364 A difference of 1.0% or 1.8% does not seem very large and may not even be statistically significant.

L391 I believe it is Chapin "III" et al. Also, this statement seems a little vague to me. Please explain what you mean.

L411 Please mention what simulation you are referring to.

L433 One reason why the modelled NEE IAV is smaller than observed may be related to the fact that your model does not represent mortality. That may be worth mentioning here as well.

L481 I assume that "models of the future" refer to models that project future changes in vegetation dynamics? Please rephrase.

L519 To decide whether a simulation is more realistic, you would need to evaluate your model output against some kind of observation-based reference data.

Tables

Table 1: Add information on how CO₂, climate, and LULCC are treated in each experiment.

Figures

Figure01 The time axis covers the period 1750-2014. Why does the caption say that the curves corresponds to the HISTORICAL and RESTART run, if the RESTART run starts in 2015?

S1b Please make larger to enhance readability.