Response Letter for "The long-term impact of transgressing planetary boundaries on biophysical atmosphere-land interactions"

Reviewer 2

The manuscript addresses some of the criticisms of the Planetary Boundary framework, especially the oversimplification of a very complex topic and the missing interactions between boundaries. Given the prevalence of the PB framework in the public debate on climate change, it provides a very valuable contribution to the discussion. I have some major concerns that should be addressed before publication.

We thank the Reviewer for the useful comments and review of our study.

My main concerns are with the validation of the model results:

- The model used in the study tries to model highly chaotic processes over very long periods of time, especially deforestation (which depends on human behavior) and vegetation itself. The authors use single model runs, a sensitivity analysis is necessary to show uncertainty of the model results. Especially as running the model does not seem to be too expensive this should not be too difficult.

Although our model is fast in comparison to other GCMs, conducting entire sensitivity studies is still unfeasible. Despite utilizing a powerful high-performance computing system, a single 800-year run demands at least four weeks, excluding the prolonged spin-up of carbon pools, which scenarios can utilize collectively. This limitation is characteristic of sophisticated Earth system models. The predicament aligns with established practices, such as the IPCC's approach of evaluating multiple models under a common protocol. Although an intercomparison would be valuable, proposing it for the 7th IPCC assessment report is more suitable, given the comprehensive nature of such an endeavor. However, this scope surpasses the intentions of an application paper on simulating planetary boundaries, where the scenario simulations serve a partially analytical purpose. A comprehensive analysis warrants a separate and dedicated paper.

We added a sentence to the computational demand of the model to the methods (l. 184-189):

"Utilizing a robust yet relatively efficient Earth system model allows us to conduct simulations spanning centuries, a departure from typical CMIP6 experiments that generally conclude in 2100. Nevertheless, even with CM2Mc-LPJmL, a single 800-year run requires a minimum of four weeks (utilizing 64 CPUs), excluding the extended spin-up period for carbon pools, which scenarios can collectively share. Consequently,

our emphasis is on no more than six distinct experiments, aiming to scrutinize the influence of various combinations of planetary boundaries on the Earth system. This approach already poses a considerable demand on computational resources."

- The model is quite novel. How well does hindcasting work? This is probably the best way to verify the accuracy of model results.

The validation and evaluation of the model has been done in the model description paper (Drüke et al. 2021). We added the following figures, adapted from Drüke et al. 2021 to the Supplement. They show an evaluation of the historic period. While there are some deviations, mostly due to the coarse spatial resolution (e.g. in the exact location of the ITZC), modeled biomass and historic temperature evolution are relatively well captured for an Earth system model with coarse resolution and dynamic vegetation. For more details refer to Drüke et al. 2021.

We added in l. 119-121 :

“POEM exhibits significant climate biases due to its coarse resolution (as shown in Fig. S5). However, it does reasonably well in capturing historical temperature trends (Fig. S6) and global biomass distribution (Fig. S7), albeit with a pronounced negative bias in Amazon rainforest biomass. For a more comprehensive evaluation of the model, refer to Drüke et al. 2021.”

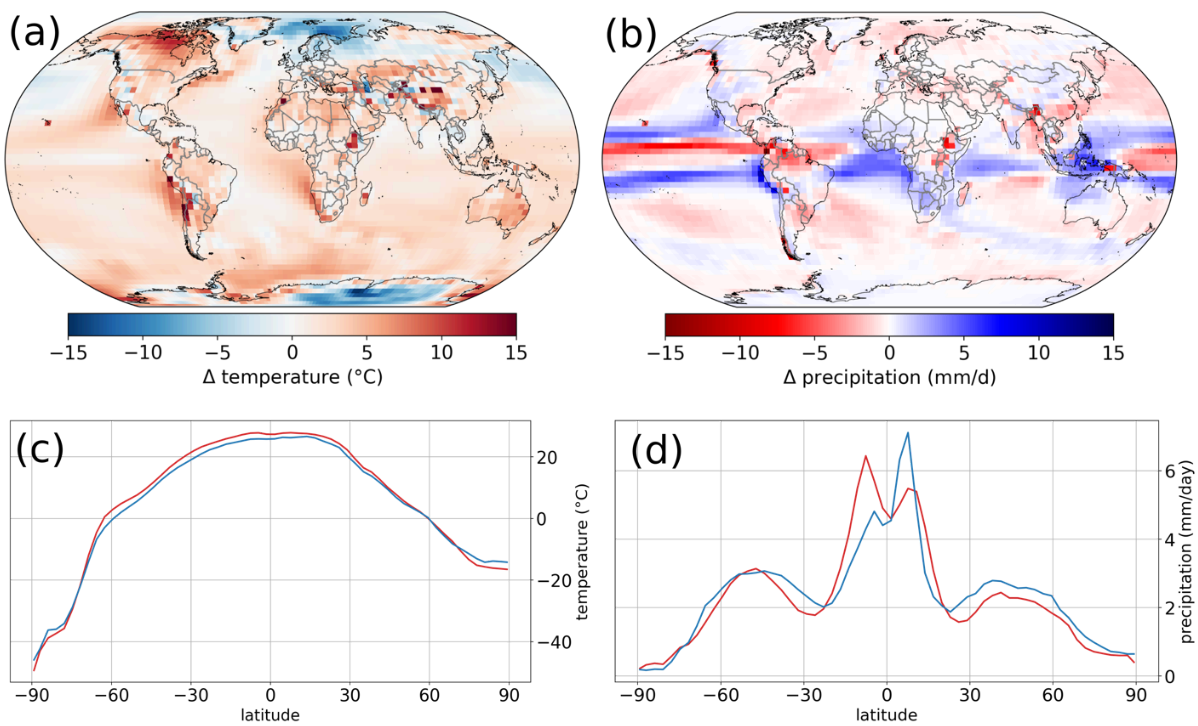

Fig. S5: (a) Global surface temperature from CM2Mc-LPJmL averaged over the period 1994–2003, (b) global precipitation from CM2Mc-LPJmL averaged over the period 1994–2003, (c) zonal mean temperature from CM2Mc-LPJmL (red line) and ERA5 data (blue line) averaged over the period 1994–2003, (d) zonal mean temperature from CM2Mc-LPJmL (red line) and ERA5 data (blue line) averaged over the period 1994–2003. Plot is analogous to Drüke et al. 2021.

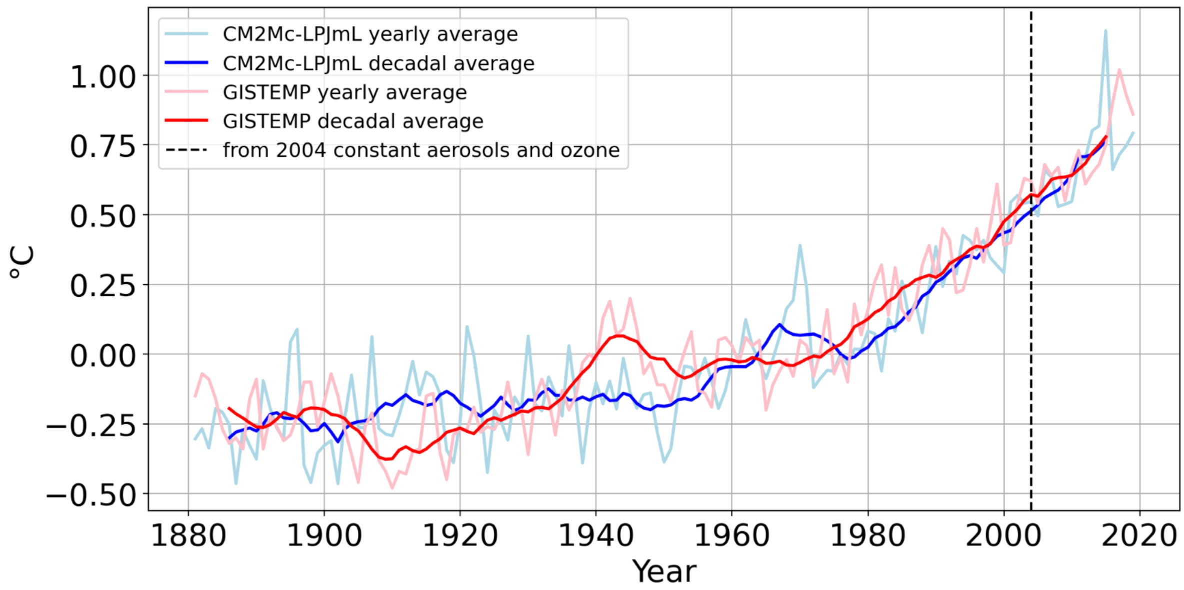

Fig. S6: Yearly and decadal global mean temperature anomaly (relative to the reference period 1951–1980) of CM2Mc-LPJmL compared to GISTEMP data from 1880–2018. Note that, from 2004 on, only greenhouse gas forcing remains, while aerosols, solar radiation and ozone are set to their corresponding 2003 values. Plot is analogous to Drüke et al. 2021.

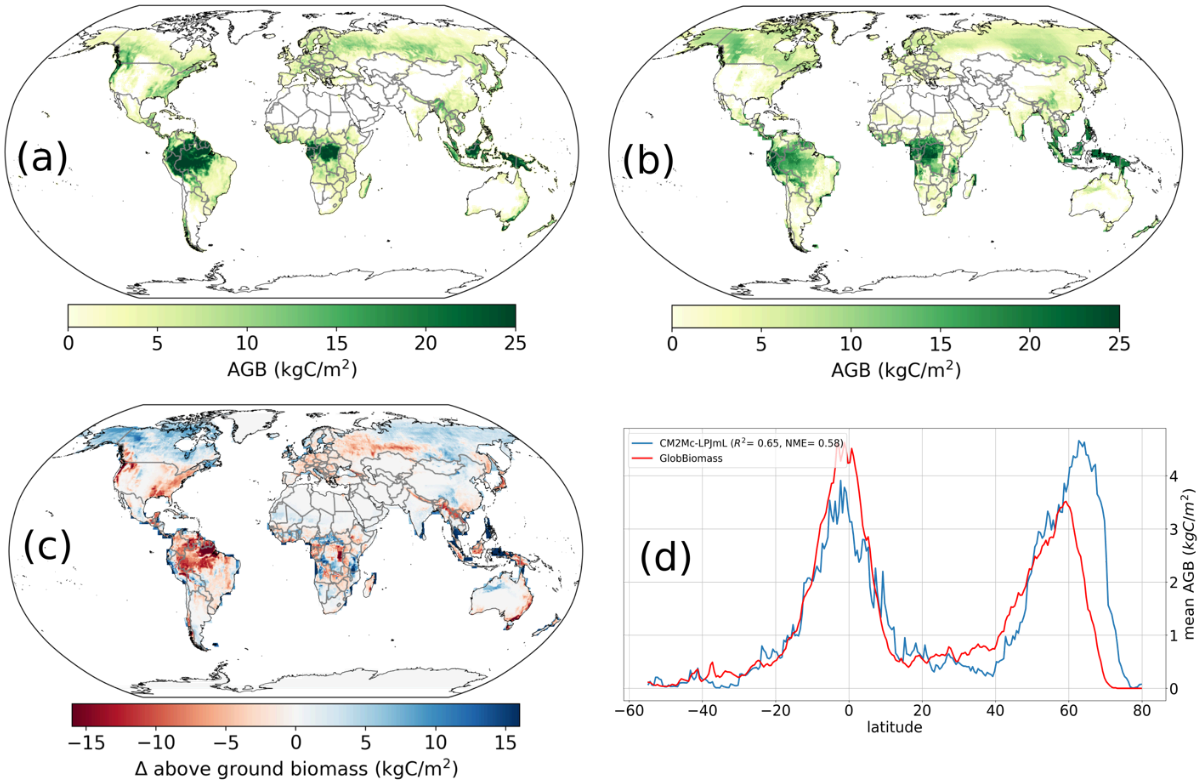

Fig. S7: (a) Mean global above-ground biomass of GlobBiomass (Santoro et al., 2020) evaluation data. (b) Mean global above-ground biomass of CM2Mc-LPJmL over the period 2006-2015. (c) Difference of the above-ground biomass between CM2Mc-LPJmL and GlobBiomass evaluation data. Blue/red colors denote an overestimation/underestimation of biomass by CM2Mc-LPJmL. (d) Latitudinal sum of above-ground biomass from CM2Mc-LPJmL (blue line, R2=0.65, NME=0.58) and GlobBiomass evaluation data (red line). Plot is analogous to Drüke et al. 2021.

- I think it is important to see the isolated effect of crossing boundaries (and I disagree with reviewer 1 here), especially because the authors put interactions between the different boundaries as one of the main aspects of your analysis. It would have been interesting to see "risky land system change" and "safe" climate change scenario, even though this is an impossible scenario. This is the only way to quantify the additional effects of crossing both boundaries.

We agree with the reviewer that simulating the full matrix of different boundary states would be interesting but as outlined earlier, our experiments actually used a lot of computational resources. As our main focus was on the analysis of long term dynamics, we decided to focus on the more probable states of the planetary boundaries and following a prioritization of computational resources. Unfortunately, doing these experiments would take up too much computational power and was therefore out of scope for this study.

I have one minor remark: In some figure the font is unreadably small

We adapted the figures and increased the font size in figures 4,6,8 and 10.