Response to Reviewer 2 comments

Review 2 has very strong comments on the content and the organization of our manuscript.

"I "recommend modifying the structure and results to better reflect the target audience. You can leave the model development and testing sections, but you need to provide additional context and analyses."
« General Comments:
The results and analyses left me unsatisfied, especially given the amount of time spent on model testing. Many findings are dismissed as beyond the scope of this paper or for future work. However, without in-depth exploration of at least some of the interesting results of the simulations, the paper feels more like a data description, which is fine, but not especially appropriate for Climate of the Past. I recommend expanding the transient simulation results and analysis, since it is the novel part of this study. There are several topics that could be explored further, such as the importance of dynamic vegetation in the transient climate response (needed for the title), the mechanisms driving multiple equilibria, and comparison with proxy reconstructions. The authors might also want to consider how do these results compare with other transient model simulation?"

Some of these comments are consistent with those of reviewer 1. We realize that the structure we adopted for this manuscript deserved us. We already provided quite a lot of in depth analyses even though we agree that the section on the transient simulation as it is appears a little bit descriptive. We propose to add a few things on the transient simulation to better discuss the response to the insolation forcing and the linkage between climate and vegetation at the regional scale. But we will keep our initial focus and use the different tests we did to highlight the context in which the simulation can be considered, in particular for future model-data comparisons. This implies that we better highlight the limits we discussed. They come from the possibility of multi-states for vegetation, model biases and caveats for model evaluation on the pre-industrial or the historical period. We thus propose to reorganize the manuscript so as to have the discussion of these points in the last section. This will allows us to better connect the different pieces and provide a more focused manuscript.

To better emphasize the results of the transient simulations we propose to restructure the manuscript as follow:

1. Introduction
2. Model and experiments
3. Simulated climate and vegetation throughout the mid to late Holocene
4. Multiple vegetation states and uncertainties
5. Conclusion

Compared to the original outline:

1. Introduction
2. Model, mid Holocene and preindustrial experiments
3. Mid-Holocene simulations with interactive vegetation
4. Simulated climate and vegetation throughout the mid to late Holocene

5. Conclusion

As stated above, the new structure is a response to the reviewer request to revisit the structure of the manuscript. The new section 2 will start for the experimental design of the transient experiment; so as to explain that the mid-Holocene is the reference period and that only a subset of simulations was run for the pre-industrial period. The discussion of the sensitivity test will be slightly refocused and redistributed in the different subsections. The construction of the MH initial state for vegetation will also be included, but not the discussion on the possibility for multi vegetation states for PI. The current section 3 on mid-Holocene simulations will thus be redistributed between section 1, and section 4 where a specific focus will be put on the multiple vegetation states for PI and the evaluation of the simulated vegetation for MH and PI using the biomisation method. We'll also emphasize what we call limits in the title. In the new section 3 on the transient simulation we'll slightly enlarge the analysis of the response to the insolation forcing and add a discussion on the climate variables over the three regions.

The different figures will be reorganized so as to reflect the new outline. It sounds difficult to reduce their number, but we'll find a way to have fewer maps with vegetation changes.

Responses to the other comments. The comments dealing with text editing will be considered if still relevant in the revised version of the manuscript. We answer only to questions or to comments considering the content.

Line 123: Can sea level actually change in the model?

The ocean model has a free surface. The average sea level evolves with the global surface water budget (evaporation – precipitation – river runoff – water flux from ice sheet). However, the numeric is not designed for sea level large sea level change. It's better to keep it small with regards to the depth of the first level (10 m). The water conservation in the coupled model is thus critical for sea-level stability and to make sure that the sea-level change in a transient experiment is indeed the result of climate changes and not of model spurious drift.

Line 143: I do not think that this is very good justification for not thoroughly testing modifications against preindustrial climate.

We do not fully understand this comment. We provide comparison for PI for a subset of simulations. The argument on computing time is only the truth and we had to adjust our strategy to the computing allocation we had. We didn't had enough computing time to run both 1000 years long simulation on MH and PI for all the tests. We started the model developments on PI and then move on MH for the final tests that are presented here that all requested long simulations.

Lines 152-153: Given the importance of the aerosol responses, why do you prescribe aerosols here? Are dust and sea-salt prescribed to PI? How might this impact climate? I do not find ": : :we also plan to run simulations with fully interactive dust and sea-salt" good justification.

Here also we do not fully understand this comment. We wrote: because we are developing the interactive version with aerosols that can be run on 6000 year long time periods. What we didn't write is that is also requires developing the full coupling with the interactive vegetation for the dust sources, and that we are here at the first step with dynamical vegetation. It is ongoing work. That

would require another 1 to 2 years. Aerosols are set to their pre-industrial values, and they are fully interactive with the radiative code in the atmosphere.

Line 163: Some of these modifications do not feel robust (e.g. the soil evaporation factor).
They are, we only show robust results here.

Lines 175-176: What is the TOA energy imbalance for these runs? This could be important since different simulations are run for different amounts of time. 0.4 Wm-2 is far from zero: : :

The imbalance is negligible ~0 (with interannual variability around it). Of course, during the adjustment phase, it is equivalent to what is shown at the surface. Part of the small offset at the surface results from small errors when estimating the heat budget from the monthly model outputs. It is not possible from the limited output we kept for the long simulations to properly reallocate the right latent heat values when we are dealing with evaporation or sublimation on surfaces with evolving sea ice or snow. This is done properly during the model run. It is thus mostly a diagnosis error rather than a model imbalance. Note however that interannual variability is of the order of 0.2-0.4 W/m^2.

Line 181: Why does this modification impact the ocean response so dramatically? I thought the hydrologic modification should only impact the land surface. Am I missing something?

Because we are in a coupled system and that energy is redistributed between land and ocean. Changes in evaporation over land affect moist static energy and its gradients.

Line 310: how well does 50% compare with other models?

We will conduct the same biomisation and evaluation against pollen data for PI outputs. This would provide a comparison of model performance in vegetation distribution under different climates. For other models, say PMIP3 outputs, it is difficult because the biomisation algorithm requires some variables (e.g. tree height) that are not usually uploaded in PMIP3. We will, however, add more discussions here about the model-data evaluation, referring to the recent work of Dallmeyer et al., (2018, CPD).

Lines 325-326: "Since surface variables adjust rapidly, this is a way to compare the rapid adjustment to insolation 326 and the additional effect due to the dynamical vegetation (not discussed here)." Why say this then? It feels like an advertisement: : :

We agree it is not needed.

Line 380: Do you mean JJAS? How do you account for the calendar changes?

We don't account for calendar change. See response to reviewer 1. The changes are limited, even though present over the last 6000 years. Here, we only discuss robust features that would emerge whatever the choice of calendar. We'll add a caution mark on the calendar in the revised version.

Line 398: Interesting: : :Worth performing spectral analysis on the variability?

Certainly, we agree, but later and not in this manuscript. It is a subject per se.

Lines 418-419: Why would this lead to an underestimate?

Even though the carbon cycle is interactive in the land surface and in the ocean, the fact that the carbon concentration is imposed in the atmosphere in the model prevents carbon feedback between the different reservoirs. Model forced in emission rather than concentrations have a larger range of response in their carbon cycle. This is why we think that in a simulation with where emissions interact with the atmospheric concentration could lead to different results. The wording here was misleading. Underestimated was there to mean not fully computed and that in a fully interactive model the results could be different. We'll revisit the sentence to better reflect what we want to say.