

Response to Reviewer 1 comments

Review one provided several important comments on the structure and the objectives of the manuscript. The major recommendation is

“I recommend to restructure the manuscript. The time-slice experiments must be embedded more strongly in the results of the transient simulation and a clear link must be established between the simulations. To reduce the number of experiments and figures, the simulations dealing with finding an appropriate initial state or discussing the differences to the PMIP3-CMIP5 model could be shifted to the Appendix. These are technically interesting but seem not to follow any scientific question. The result section of the transient simulations should be extended and more specified. In addition, research questions and aims of the study should be worked out to give the results a clear framework.”

We agree while reading these remarks that the original outline of the manuscript doesn't put enough emphasis on the transient simulation and that it would be better to construct the outline of the paper so as to better echo the title. It is important for us to keep the discussion of the different sensitivity tests. This knowledge is needed to properly analyze the results of the transient simulation and to know what we can or cannot expect from it. We will add a few results on the transient simulations. But we'll keep most of the content as it is. 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

The new structure is a response to the reviewer comment to provide a clear framework for the results. The new section 2 will start from the experimental design of the transient experiment; so as to explain that the mid-Holocene is the reference period and only a subset of simulations were run for the pre-industrial period. The discussion of the sensitivity tests will be slightly refocused and redistributed in the different subsections. The discussion on the MH initial vegetation state will be included, but not the discussion on the multi vegetation states for the PI vegetation. 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. This is a way to discuss 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 at the regional scales.

The different figures will be reorganized so as to reflect the new outline. It sounds difficult to reduce the number, but we'll find a way to have fewer maps with vegetation changes. It requires some work, but it should be easily done, thanks to the way we organized the model outputs needed to prepare this manuscript.

We also would like to thank reviewer 1 for the list of minor comments that are useful to improve the manuscript.

All the editing comments have been taken into account and already added in the text before any change is made. We provide below some responses for the other comments

L130: what do you mean with 'transient late Holocene simulation'

The last 6000 years (I.E from -6000 BP to 0k = 1950 for insolation). This will be stated more clearly in the text

L243: Please explain the metrics in more detail (e.g. in the Appendix) because the metric package may be unknown to the readers

A paragraph will be added in the appendix to better explain what is computed

L314: The heading of chapter 3 is: 'mid-Holocene simulations' so why is there a section dealing with pre-industrial climate?

We hope it will be less misleading in the new outline. The point is to know how good is the model quite early in the text. In the new version we decided to evaluate the "climate" in section 2 and have the discussion on "vegetation" in section 4. This should better insist on the fact that we have active dynamical vegetation in this simulation and that considering climate or vegetation evaluation can lead to different conclusions on the realism of the simulation depending on the way the evaluation is done.

L319-320: I do not understand what is meant by 'vegetation biases' in this context. When vegetation is interactive, the calculated vegetation distribution can be biased, but how does this bias impact the representation of the simulated vegetation? Please clarify.

We only have in mind the biases coming from climate-vegetation feedbacks that amplify the known bias of the model when dynamical vegetation is switch off. We are not in a position where we can tell how the bias in the vegetation model affects the full coupled system.

L337: What do you mean by this? that the differences in PI simulations are of similar magnitude as the differences between PI and MH ?

Since the vegetation map are similar in Vmap and Vnone for MH, the difference in PI vegetation between Pi-Vmap and PI-Vnone explains the difference in MH-PI vegetation calculated using the Vmap simulation or the Vnone simulation. We'll revisit the way we discuss it.

L383: ': : : follow the long term insolation changes in each hemisphere: : : ' What about SH Winter? Please be more precise.

We'll add a discussion on this point, but focusing on the seasonal cycle and the seasonality of the insolation forcing. For the northern regions and the southern hemisphere, part of the answer is in the ocean heat storage and the other part is in the sea-ice and snow cover.

L409: It seems as if the tree fraction follows the summer insolation change. Please specify and explain. What about the annual mean changes in temperature, precipitation and insolation?

We will add a short discussion on temperature and precipitation, but for the 3 regions we consider later in the text, considering, min, max and annual mean monthly temperatures and precipitations as well as sea-ice and snow cover for the region north of 60°N and Eurasia.

L456: What do you mean by 'rapid changes' and if these 'deserve attention' why don't you investigate them in this study?

We should have included these remarks in the conclusion. It is out of the scope of this paper. So we'll refocus the text.

L476-477: Is there a possibility to figure out the reasons for having different PI climate-vegetation changes?

We provide all what we know and the possible caveat in the manuscript. Going further requires a new study and certainly another 1 to 2 years to do it properly with ensemble sensitivity tests. We already checked all what we could check in the last 2 years about it. This is also why it is important for us to show it and discuss it in the manuscript. It can be "by chance" or linked to amplification of small differences in the initial state under modern conditions.

-L581: Isn't it originally the method of Prentice et al. 2011? What is different to the method of Zhu et al. 2018?

Yes, the algorithm follows Prentice et al. (2011), with thresholds prescribed as in Zhu et al. (2018). We also tested the different threshold values reported in Figure A2.

We will revise this sentence as: "To convert the modelled PFTs by ORCHIDEE into mega BIOMES, we use the algorithm proposed by Prentice et al. (2011). Figure A2a shows the different threshold values tested in this algorithm, with the black numbers corresponding to the default values used to produce Figure 7 in the main text."

-L590: It is not obvious why the GDD limit of 500_C is being tested, are these values realistic? I guess ORCHIDEE also uses a GDD limit of > 350_C for the existence of boreal trees vs. tundra (GDD5<350_C). A biomisation using a GDD limit of 500_C thus may not represent the vegetation simulated by the model, because it suggest tundra in regions that are suited for forests.

The threshold of 500 °C days is tested because it was used in Joos et al. (2004) to convert LPJ-simulated PFT fractions into biome types. ORCHIDEE intrinsically does not use a simple GDD limit to constrain the existence of boreal tree PFTs. GDD thresholds are only used in the phenology module to determine the onset time of leaves,

while their values are PFT-specific and are also modulated by the dormancy period, which varies for the same PFT located in different grid cells (see more details in Krinner et al., 2005). By influencing leaf onset, GDD values impact photosynthesis and growth of the PFT, and then indirectly affect establish/mortality rates and finally abundance of this PFT in ORCHIDEE. The biomisaztion algorithm is just a post-processing of the fractional PFT outputs of ORCHIDEE, with some broad-scale empirical thresholds. Therefore, we do not think testing a value of 500 °C days here would be “incompatible” to ORCHIDEE-simulated vegetation.

*-L598-599: Should we now reconsider the choice of bioclimatic limits in the DGVMs?
- What about data availability?*

As mentioned above, GDD is not a direct bioclimatic limit inside ORCHIDEE. Furthermore, although changing the GDD limit to 500 °C days improves the metric for tundra in Figure A2b, we should keep in mind that (1) any bias in simulated temperature will also affect the biomisaztion result and thus the “correctness” compared with the pollen data; and (2) the expansion of tundra over woodland in the case of “GDD=500” compared to “Default” might actually degrade the biome distribution, which cannot be reflected in the “correctness” metric because of limited pollen data in middle Siberia (this is why we mentioned data availability).

-Fig.4: Why are MH_Vnone and MH_Vmap so different?

We are not sure we fully understand the question. These simulations start with very different initial state for the land surface model. So it reflects different adjustment time, and the curve show they converge to the same solution. So we would rather say that they are very similar and not different.

Fig.7: It should be explained, why there is Savanna in the northern latitudes. In my print, the pink and orange color is not really distinguishable. Please state, why there is no grassland in North Africa in your simulation

As shown in Figure A2a, “savanna and dry woodland” is defined if the foliage projective cover (a combination of simulated fractional coverage and leaf area index) is high but average tree height is not enough. Since tree height is mainly determined by woody biomass in ORCHIDEE, we speculate that a potential underestimation of tree biomass in the model might lead to the replacement of boreal forests with woodlands in the high latitudes. This could be because of bias in climate and/or bias in ORCHIDEE in terms of photosynthesis or carbon allocation scheme. We will add these discussions in the corresponding text.

For North Africa, the model simulates desert instead of grasslands. This is mainly because of amplification by the climate-vegetation feedback of the underestimated precipitation in this region.

We will change the colors to make them more distinguishable in the revised manuscript.

-Fig.9: Maybe this figure could be moved to the Appendix

We will keep this figure in the text and add to it a panel with the seasonal change in incoming solar radiation at TOA in both hemispheres. Showing the forcing we use in the simulation over the 6000 years is important for the discussion. We’ll also better emphasize in the text the result of the last period.

-Fig.10: When looking into palaeo-seasons, one always faces the problem of different calendars. The months NDJF or JJA differ in length between mid-Holocene and PI. It should at least be mentioned in the text and in the caption, that this 'problem' exists and is not considered, neither by the model nor in the analysis. But this problem may change the trends discussed here!

In practice the effect of calendar over the mid Holocene is small. Jousaume and Braconnot 1997 show it is 5 days at most for the difference in the date of the Autumnal equinox when March 21 is prescribed as the reference date for the vernal equinox in all simulations. We are averaging on long time and show the long term trends, discussing only the significant results. The larger analysis biases resulting from the calendar are found in autumn. We do not discuss this particular season. Our conclusions, given what we are doing here will not be altered by the calendar effect. But we recognize that we need to keep this in mind. For other periods, when eccentricity is larger, this would not be the case.

-Fig.11: it is 'Northern Hemisphere'. It would also be interesting, how the simulated tree cover and bare-soil fractions at the end of the simulation compare to modern estimates on (natural) tree cover. How large is the underestimation of forest in the high northern latitudes by the model?

We agree it would be interesting, but we are also concerned that because we do not have land use in the simulations. Land use has an impact at regional scale. However we also know, and this is shown in the MH biome comparisons, that the differences between the simulated vegetation and the real world are larger than differences that would come from land use. Since we decided now to add a biome comparison for PI using pollen data for 0k, we'll consider this remark in the revision. We'll also reinforce the discussion and questions about the evaluation of the vegetation we simulate out of such transient simulation.

Fig.12: What causes the strong peak around 4.8ka ?

This event comes from internal noise and/or compound variability events, superimposed on the long term trend induced by the insolation forcing. There is no obvious cause. We checked that it doesn't come from an artificial computing failure when running the simulation.