Referee #2:

The authors have addressed most of the comments raised in previous round in the revised version of the manuscript. There are, however, a few concerns that require the authors to clarify before a final publication:

(1) For all the climate variables used by BIOME4, the authors use "absolute difference" between the climate of each time-slice and the pre-industrial climate, and add the "absolute difference" to the reference present-day climate to generate climate data for each time-slice. This method works well for some variables (e.g., temperature), but can be problematic for other variables (e.g., precipitation). Precipitation can be minus values, if "absolute difference" is too negative. Have the authors checked if there are any problems with adding "absolute difference" to different climate variables? How sensitive are the BIOME4 model results to the ways of adding "anomalies"? More explanations is needed to confirm if the "absolute difference" approach is valid for all variables.

A: This is indeed not specified in the manuscript. Using this anomaly approach can lead to negative values for precipitation and sunshine. In case negative values occurred, we set them to zero.

We added this information to the manuscript and now write: „Negative values in precipitation or sunshine resulting from a too large negative difference of the variables between 6k and 0k, compensating the present-day values, have been set to zero."

This method is a common method used in biome-model studies, see e.g. Wohlfahrt, J., Harrison, S.P., Braconnot, P. et al. Clim Dyn (2008) 31: 871. doi:10.1007/s00382-008-0415-5. or Wohlfahrt, J., Harrison, S.P. & Braconnot, P. Climate Dynamics (2004) 22: 223. doi:10.1007/s00382-003-0379-4.

(2) Fig. 1: Please add "(GPCP dataset for precipitation, ERA40 for wind vectors)" instead of "(based on reanalysis data)".

A: done.

(3) Fig. 11: Time axis of the figure is still not changed. "surrounding of both sides" should be "surrounding of both sites"? The region (106°E-110°E, 69-73°N) selected for the southern Taymyr record does not surround the locality of the record (72.4°N, 102.29°E). The reason for this is not clear to me. It is said that "the modern taiga-tundra margin is shifted westward in the modern biome simulation". According to this, the region selected should also be shifted westward. But the authors selected a region (106°E-110°E, 69-73°N), which in fact, is to the east of the proxy locality. Why?

A: Thank you, this is indeed a mistake, the Taiga-tundra margin is shifted northward. We simply take the nearest grid-box around the lake showing tundra vegetation in the modern climate. We change this information to: „The latter region had to be adjusted to the modern taiga-tundra margin that is shifted northward in the modern biome simulation due to an overestimation of forest cover in the boreal region. For comparison, the nearest grid-box around the lake showing tundra vegetation in modern climate has been taken."

Regarding the time axis, we have changed the labeling, but unfortunately put an old version of the figure into the revised manuscript. We now changed it.

Referee #3
Review for "Biome changes in Asia since the mid-Holocene – an analysis of different transient Earth system model simulations"

This manuscript describes the vegetation changes in the Monsoon-Westerly (forest-steppe-desert) transition zone and high northern latitude (taiga-tundra) transition zone in Asia since the mid-Holocene using several different transient simulations and a slightly modified BIOME4 vegetation model. To validate the simulated vegetation changes in the two areas, the authors also try to implement model-data comparison.

The analyses applied to this study are not new (in particular vegetation changes between 0k and 6ka), and thus there are still some problems that should be resolved for data-model comparison with regard to paleo-vegetation changes. Because the paleoclimate and paleovegetation community already know the arguments about data-model comparison in this manuscript, the authors might have to propose a new approach. Anyway, the authors did a great job in analyzing so much data and in producing very nice figures.

The topic is really interesting, but I had several problems here.

Specific comments

L95-98. Although I understand the three research aims in this study, I am not sure that the authors use reasonable approaches for achieving the goals. Data-model comparison approach in this study is not new, and thus we might not improve our knowledge about uncertainties of simulated vegetation changes?
A: The major aim of this study is to provide a range of possible biome changes in East Asia during the last 6000 years. We agree with the Referee that the general methods used in this study are not new, but – at least to our knowledge - a multi-model analysis has not been undertaken for Asia so far, nor transient simulations have been used before. Using transient simulations enables us to analyze trends in biome change. Furthermore, we improved the BIOME4 model for the Asian region.

It seems that the authors mainly use multi-model ensemble approach for the transient vegetation change analyses (Figure 6 - 11), but the authors show the results from the several simulations including the ensemble mean for mid-Holocene vegetation changes (Figure 3 and 5). If possible, you show the results from the ensemble model mean for general trends in main text and each model result in the Appendix because you do not discuss the climate and vegetation changes and their causes among the models thoroughly.
A: We keep the figures 3 and 5 as they are, because we want to provide a range of possible biome distributions for the mid-Holocene time-slice and want to stress that the major vegetation changes and the climate factors yielding the vegetation change are robust, i.e. similar in all models. Mid-Holocene is a key time-slice in the PMIP analysis and, therefore, of special interest for model evaluation and the reconstruction community as synthesis of reconstruction and many climate simulations exist for comparison.
The causes of the differences between the models can not be answered thoroughly without further model experiments.

L105. "absolute minimum temperature" for "annual minimum temperature"?
A: done.

L112-113. The authors need many PFTs and thus biomes for describing the diverse taxa found in Asia. On the other hand, the authors grouped the BIOME4 28 biomes into 12 mega-biomes for the

analysis. It sounds that your argument and your approach is inconsistent here. Moreover, the authors discuss the vegetation margin analysis using only "desert", "steppe", "forest" and "tundra" later. If you use these limited biomes, you can use other process-based vegetation models such as LPJ-GUESS. Your target areas for the vegetation changes in Asia are limited (only transition zones), and the target biomes are also limited.

A: Showing and discussing the change in all 28 biomes used in BIOME4 would result in a very long manuscript. We therefore decided to follow the approach of Harrison and Prentice (2003) to group the biomes in mega-biomes. This has also the advantage that one eliminates indication of small model biases, for instance, when the model simulates biome types which are not the same, but only similar to the observed ones (e.g.cold instead of cool mixed forest). Using only macro-biomes in the 'transient' biome change is also a matter of representation and reducing model biases. But nevertheless, the advantage of using BIOME4 is that all individual biomes are calculated by the model and their climatic tolerance and other requirements are included in the calculation. They are grouped afterwards. Therefore, BIOME4 can respond more specific to the climate forcing than other models, that only use desert, forest and shrubs as vegetation types, and we wanted to calculate biome distributions as pollen-based reconstructions are often assigned to biomes.

(Harrison, S. P., and Prentice, I. C. (2003). Climate and CO2 controls on global vegetation distribution at the last glacial maximum: analysis based on palaeovegetation data, biome modelling and palaeoclimate simulations. Global Change Biology 9, 983-1004.)

L120-123. Although I checked the Table 1 for bioclimatic limits for the model, I do not understand how the authors used the altitude information from ETOPO5. If possible, you should describe how to use the information? Moreover, do we need to change the surface air pressure for running BIOME4 at high altitude grid points?

A: The bioclimatic limits are switched to the „altitudinal vegetation constraints" in grid-boxes with orography higher than 2000m. We added this information in the caption of Table 1: „The limits in the brackets represent the climatic tolerance on mountains (altitudinal vegetation constraints) to which the model switches in grid-boxes with orography exceeding 2000 m."
Regarding the surface air pressure: we looked into the model code and could not find any equation in which the surface air pressure is used.

L134-135. What is 0k and 6k here? You need to define these words before this sentence or here. Otherwise, you can delete the texts after "(Fig. A2)".
A: We changed 0k to pre-industrial and 6k to mid-Holocene.

L138-141. In the previous studies with BIOME4 for simulating paleovegetation, they used CLIMATE2.2 climate data. Why do use CRU TS3.10, not CLIMATE2.2 in this study? If you need higher-resolution of modern climatology as reference data, why do not you use CRU CL2.0 (New et al., 2002)? The data includes the necessary climate variables for BIOME4 and elevation information.
A: We took the newest climate dataset available for us.

L 144-145. To simplify/ignore the impact atmospheric CO2 concentration on vegetation, the authors might use 280 ppm for 0k simulation. If so, you can delete this sentence and the related figure because you do not need to discuss the atmospheric CO2 effect to vegetation in this study.
A: We included this information and the figure at the suggestion of Referee #1 in the first revision. We keep this in the manuscript.

L 151-152. I do not understand what "the main biome distribution" is. Do you mean that the model sufficiently simulates the large-scale (or continental scale) biome distribution in Asia.
A: This is indeed imprecise. We change it to 'large-scale biome distribution'.

L156-168. These sentences describe results and your arguments. Thus you should not put these sentences here.

A: The entire chapter 2.2 deals with model evaluation for pre-industrial climate, it contains methods as well as the description and discussion of the reference simulation. We do not interpret this as results of our study, but rather as detailed description of the model. So we keep these sentences in this chapter.

L179. Remove "of the year"? This absolute minimum temperature is based on an annual data or climatology (e.g. 1960-2000)?

A: done.

L225. "CRU TS3.10" for "CRUTS3.10"

A: Thank you, we changed it.

L229. "(Harrison et al., 1998)" for "(Harrison, 1998)"

A: Thank you. Done.

L230-231.According to Harrison et al. (1998), the anomaly approach has two advantages; 1) reducing the effects of systematic model biases and 2) capturing some of the locale-scale spatial pattern because of terrestrial geography. The authors already describe these advantages in this part, and I understand them. But I do not understand the sentence "We are however aware of… meso- and large scale" Do you need this sentence and the next sentence in this section? I mean, you can describe the methodology here.

A: We rephrase these sentences more precisely and now write: „We are however aware of the simplifications inherent to this approach in interpolating coarsely resolved GCM output (resolution approx. 3.75° and coarser) onto higher target grids (here 0.5°) without taking into account potentially important factors that lead to local variations in climate, such as changes in variability and feedbacks from the local to the meso- and large scale."

L243-244. For comparison, the authors choose one high quality pollen record at each key transition area. But the model output is based on 500-year interval data. Thus, if possible, the authors should use relatively low temporal resolution pollen data. I think one data at each zone is very small number, they might not validate your results well.

A: We agree that higher data coverage would support our results better than using only one record. We have discussed this comment already in the first revision (c.f. author response to Referee #3). High quality data is quite few. Therefore, we decided to select from all (for us) available records in the target-regions the record, that shows the best quality with respect to dating and data and that is additionally representative for the regional vegetation changes, as inferred from comparison to literature results. The general vegetation trend indicated by the Dahai record is in line with other records in North central China (cf. Zhao et al, 2009) and the 13-CH-12 record agrees with the other records from the northern Siberian tree line area.

L260-272. As mentioned earlier, the authors can describe results from only ensemble model mean because "the vegetation change is small and similar for all models". You can move the other models output into the Appendix. Moreover, You can combine two figures (Figure 2 and Figure 3) into one because you do not need to show "CRU TS3.1 0k" twice.

A: We decided to keep the presentation and discussion of the mid-Holocene (6k) biome distributions simulated by the individual models in the main text, because one aim of the study is to show a range of possible mid-Holocene biome distributions. This is usually the time-slice compared with reconstructions as high-quality synthesis of reconstructions exist for 6k. We also keep the CRU

TS3.1 0k map in both plots, to facilitate the comparison with the reference map and the 6k biome changes.

L282. Delete "East"? Your target area is the entire Asia, not East Asia right?
A: We changed this to Eastern Asia, to be consistent with the other sections.

L283-289. This paragraph (except the first sentence) describes a method for a sensitivity study, not results. Thus, you should move this content to the "Methods" section.
A: We agree, this paragraph rather describes the method of the sensitivity experiment and not results, we move this paragraph to the method section (L238-243) and shorten the information in this paragraph to: „The Holocene changes in bioclimate are discussed in the Appendix (Fig. B1). To assess the climate variables being responsible for the biome shifts in the model since the mid-Holocene, climate variables of the pre-industrial input dataset are gradually replaced by the respective variable of the simulated mid-Holocene climate."

L288. "CRU TS3.10" for "CRUTS3.10"
A: done.

L288. "absolute minimum temperature" for "absolute annual minimum temperature"
A: done.

L290-298. This result is really interesting for me, and it is easy to know the causes of vegetation changes over each transition during the mid-Holocene. Do you make the map from ENSMEAN for Fig.5? Moreover, this information is directly concerned with vegetation changes between 0k and 6k, and thus you can put this information into 3.1.
A: We have not prepared a map based on the ensemble mean climate since the results of the individual simulations are very similar.
We decided to discuss the climate factors in a separate section (3.2).

L291. Your sensitivity experiment cannot deduce "the prolongation of the warmer season" because these climate variables does not have temporal concept probably. We might deduce warmer during the growing season in the northern latitudes (from the table 1: tundra vs. boreal forest and Fig. 5).
A: We are not sure if we understand this comment correctly, we infer the prolongation of the warm-season from the increase in growing degree days.

L301. "46%" in the text, but "45%" in the table.
A:Thank you very much, it is 45%.

L313-. Although "geographical distances are given in degrees of latitude or longitude in the following, for simplification, only ° is used", the authors use "100 km" here. Should you use the unit ° here? The same issues are on L507, L520, and L521.
A: We agree, this is misleading. We did not want to exclude other units, so we change this sentence to: „(for simplification, only ° is used for geographical distances given in degrees of latitude or longitude in the following)."

L375. Is it possible to compare between affinity scores (reconstruction) and fractional area cover (%) of biomes. I hope you understand the concept of biomization.
A: We compare the reconstructions and the simulated biome fractions only semi-quantitatively, i.e. both methods indicate the dominant biome type but tentatively also changes in the relative importance of sub-dominant biome types. But of course affinity scores provide no information about past coverage changes.

Discussin. The authors focus on the two point of discussion. I think you have several arguments in the methods and results sections. Moreover, to support your argument/discussion, you should use other references effectively.
A: We agree, that we discussed the model and the reference simulation in the method section, but we interpret this as evaluation of the model and not as discussion of our results. Therefore, we keep the structure.
We also agree, that there is an imbalance between citations of own studies and other references in chapter 4.1. We add references in the discussion about the impact of differences in interactive model components to the Holocene climate change:
L433-435: „Previous modeling studies indicate that land-surface feedbacks with the atmosphere could have enhanced the Holocene precipitation change induced by orbital forcing (e.g. Wang, 1999; Texier et al., 2000, Diffenbaugh and Sloan, 2002; Li et al., 2009)."

L398. I do not know whether your approach is correct or not. At least, using several transient climate model output, you run the BIOME4, and then you get the simulated biome over the last 6000 years. But, I do not think this small number of output does not show a range of possible biome changes in Asia.
A: We used all (for us) available transient climate simulations, so far, this analysis of transient biome change is all that can be provided by the climate modeling community.

L398-435. I do not understand your discussion (4.1) well. Does your discussion directly have a concoction with your result (which ones)?
A: The simulated biome distributions depend on the climate input. Differences in the simulated climate between the individual simulations lead to differences in the biome distributions. Therefore, we assess the performance of the models in simulating the Asian climate and discuss the factors (differences in model setup) that may lead to differences in the climate. We further explain this in the revised manuscript:
L403-406: „The outcome of this analysis and its reliability depends directly on the capability of the models to simulate the mid-Holocene to pre-industrial climate change. Biases in the models and differences in the simulated climate between the individual models lead to errors and differences in the simulated biome distributions. "

L442-445. I understand your argument here, but you need a limited number of biomes in this study. Moreover, Basil Davis (Université de Lausanne) has a funded project, HORNET (Holocene pollen based climate reconstruction for the Northern Hemisphere extra-tropics) and he makes biomised biome data over the Holocene, so far as I know. Thus, your argument will be solved soon, and you might use his database probably.
A: Thank you very much for the information, we are also in contact with the PAGES LandCover6k initiative and work on the biomisation of pollen-based vegetation reconstructions in several projects, but this is ongoing work.

Table 4, 5, and 6. Do you need to show the all model results here? If you describe the general/common/robust pattern, you can show only 0k and 6k (minimum, ENMEAN, maximum) values without model names.
A: We prefer to show all model results in these tables, as we discuss the positions of the desert-steppe, steppe-forest and taiga-tundra border simulated by the individual models in the Appendix and do not want to show the table twice.

Figure captions. "CRU TS3.10" for "CRUTS3.10" and "BIOME4" for "Biome4"
A: Thank you! We changed it accordingly.

Figure 11. Please put (a), (b), and unit of model output on the figure.
A: We put a) and b) on the figure and added to the caption: „The reconstructions are given in fractional coverage per area [%]."