We thank the reviewer for his/her comments which we found most useful and have added much to the discussion. Below, we reply to each of the comments which are in red font.

General comments

1. The authors have included the EPlio1 experiment in the present study to compare with the PlioMIP2 core experiment. However, it lacks broadly discussion about the difference between these two experiments with regards to the SAT, SST, precipitation and sea ice. The Eplio1 differs from the PlioMIP2 not only in the topography, which is mostly mentioned in the results, but also in the land-sea mask, the lakes in the Africa as well as the soil types. I suggest the authors could pay more attention with regards to these aspects.

We have expanded our discussion on the differences between Eoi^{400} and Eplio1 at the end of each subsection and included a supplementary figure to show the distribution of the differences. Any differences in the climate resulting from changes in soil type or the inclusion of Pliocene lakes in Africa appear to be very subtle, and an analysis would probably be better served with sensitivity experiments targeting these conditions individually.

2. Section 4.6: The Meridional heat transport have been studied in many PlioMIP2 papers (Li et al.,2020, Chandan and Peltier et al.,2017, Tan et al.,2020, Feng et al.,2020 etc). I suggest the authors can compare the results with the others. According to Figure 13, it is worth noting that the import change appears over the low latitudes which is different from the aforementioned studies. By the way, Figure 13 is not easy to read, I suggest to put all the absolute values in one plot.

As suggested, we have compared the meridional heat transport in our model to those with other PlioMIP2 models. This has been done for both the Atlantic Ocean and the total globe (atmosphere plus ocean), wherever possible, although in some studies, the anomalies are not shown explicitly. We have also made a note of the fact that the largest anomalies in the total meridional heat transport occur at the low latitudes, something that is not seen in other studies. We have put all the absolute values in both Figures 12(a) and 13(a). However, because the magnitudes of the differences for the total transport are small compared to the absolute values, it is not very easy to distinguish between the various experiments in Figure 13(a), particularly the non-core experiments.

3. Section 4.7: I do not understand why the authors compare the E280 with the proxy data reconstructed for the mid-Pliocene and discuss the performance of the data-model fit of the Pliocene experiment and the Pliocene experiment together (line 316). Section 4.8: I suggest the authors could also compare the model outputs with the PRISM4 data (Foley and Dowsett, 2019) which is more specific for the PlioMIP phase 2.

Both E^{280} and Eoi^{400} (in addition to the other Pliocene experiments) need to be shown in the comparison because the proxy data refer to absolute values of the temperature, and not the mid-Pliocene – Pre-Industrial anomalies. Without plotting the E^{280} values, it is not possible to know whether the inclusion of Pliocene boundary conditions leads to better agreement with the proxy data. In our original manuscript, we compared results from

Eoi^{400} with PRISM4 data and the results were plotted in Supplementary Figure 2 which we will now move to the main part of the manuscript. We have now expanded the text to include a comparison with the Pliocene experiments with other CO_2 levels, except Eplio1. We find that, for all CO_2 levels used in the study, not only is there still an underestimation of the warming in the northern North Atlantic Ocean and Greenland and Norwegian Seas, but also in the South Atlantic Ocean, near southern Africa.

Specific comments

1. Line 145: Confused. Please specify which 4 experiments are further integrated for another 1000 years and why they need to be integrated for another 1000 years

That sentence (and the whole of section 3.2) refers to the experiments, Eoi^{280}, Eoi^{350}, Eoi^{400} and Eoi^{450}. They need to be integrated for another 1000 years because, up to that point, the values of CO_2 and the other greenhouse gases are slightly different to those specified in the PlioMIP2 protocols (Haywood et al, 2016). The values then follow the protocol for a 1000-year model integration, as plotted on the right-hand side of figure 2. This issue is related to the point directly below.

2. Line 155: Why are the greenhouse gases not changed to be PlioMIP level in the first 3000 years run?

Our original intention was to remain consistent with our previous study for the first phase of PlioMIP which did not explicitly state the Pre-Industrial greenhouse gas levels. This was left to each modelling group. For example, our Pre-Industrial CO_2 was set to approximately 285ppm, and double CO_2 was based on this value. All 8 experiments based on these greenhouse gas levels are shown on the left-hand side of figure 2. However, it soon became clear that all other modelling groups were using levels specified in PlioMIP2. In order to be consistent with them and facilitate model intercomparison, we decided to continue these experiments with the exact PlioMIP2 greenhouse gas levels for another 1000 years. This was preferable to restarting from scratch because the first set of experiments (on the left-hand side of figure 2) had already reached a state of near-equilibrium.

3. Line 167: In my opinion, figure 3a does not approve "Temperature increase fairly uniformly", In figure 3a, the warming amplitude over high latitudes is stronger than the low latitudes, the land warming is larger than the ocean broadly, and some extreme warming regions are found in the Barents sea and around coastal regions of the Antarctic.

Yes, we agree. In comparison to the Pliocene experiments, temperature changes in E^{400} are not as extreme, but, in hindsight, we should have avoided the word 'uniform'. The text has now been changed. A description of the greater warming over land, in the Barents Sea and in some coastal parts of Antarctica was already given at the beginning of section 4.1.

4. Line 184: Compared to the PlioMIP1, PlioMIP2 also modify the Bering strait and Northern Canadian Archipelago regions, do they have any impacts on this different warming?

Cooler water in the Arctic Ocean is prevented from flowing into the Labrador Sea via the Northern Canadian Archipelago regions in PlioMIP2. This leads to warmer SST in the Labrador Sea. However, surface air temperature over the Labrador Sea remain the same or actually increases slightly because of the influence from the higher elevation over southern Greenland and North America in PlioMIP2. This higher elevation has a large influence on the northern hemisphere surface air temperatures. To better quantify the impact of closing the straits on the difference in warming would require a separate PlioMIP2 experiment with open straits.

5. Line 196: "near-uniform increase outside the polar regions" is not appropriate, according to Figure 4, I suggest to be "near-uniform increase outside the low latitudes (30N-30S)".

We have changed this part to 'a small but gradual and near-linear increase in temperature anomaly starting from the southern mid-latitudes to the northern mid-latitudes'.

6. Line 194: What is the "bias" here?

We have removed the word 'bias' and referred to the larger temperature anomaly north of the equator compared to south of the equator.

7. Line 205: What is the reason for the seasonal change of the arctic region, is the sea ice fraction change responsible for that?

Yes, the seasonal changes in the SAT of the Arctic region are related to the changes in the sea ice. In the summer, there is very little to no sea ice in the Arctic in Eoi^{400} and so the ocean warms up more from incoming insolation. The SST anomaly is at its maximum during the summer. As the summer ends, heat from the ocean is released into the atmosphere. Since there is practically no summer sea ice in Eoi^{400}, more heat can be released, explaining the greater surface air temperature anomaly in the Arctic during September to November. This was also seen in the study by Zheng et al (2019):
Zheng, J., Zhang, Q., Li, Q., Zhang, Q. and Cai, M.: Contribution of sea ice albedo and insulation effects to Arctic amplification in the EC-Earth Pliocene simulation, Clim. Past, 15, 291–305, https://doi.org/10.5194/cp-15-291-2019, 2019.

8. Line 221: What is the potential reason for the SST anomalies for the other regions, e.g. the Labrador sea, the North Atlantic Ocean?

The SST in the Labrador Sea is warmer in Eoi^{400} than in Eplio1 because the latter case had open Canadian Arctic Archipelago Straits (CAAS), and cooler water flowed from the Arctic Ocean to the Labrador Sea via the CAAS. Warmer SST in most parts of the North Atlantic and cooler SST further north, near Iceland are also seen in previous sensitivity experiments (not part of the present study) whereby we closed the CAAS and Bering

Strait in a Pre-Industrial scenario. Changes in ocean circulation led to these changes in SST, similar to those seen in similar sensitivity experiments performed in Otto-Bliesner et al. (2017).