Review comment on cp-2021-101
Anonymous Referee #2

Referee comment on "The long-standing dilemma of European summer temperatures at the Mid-Holocene and other considerations on learning from the past for the future using a regional climate model" by Emmanuele Russo et al., Clim. Past Discuss., https://doi.org/10.5194/cp-2021-101-RC2, 2021

Review Russo et al. “The long-standing dilemma of European summer temperatures at the Mid-Holocene and other considerations on learning from the past for the future using a regional climate model”

General comments

This paper presents the results of a set of simulations performed with the COSMO-CLM limited area model, with a focus on Europe and the Mid-Holocene (MH) climate. The main objective is to shed light on the question why climate models produce warmer summer conditions in Southern Europe relative to the preindustrial (PI), in contrast with pollen-based reconstructions that suggest cooler conditions. This is a highly relevant issue in (paleo)climatology, as it is important to know to what extent climate models can reproduce climate’s sensitivity to a change in radiative forcings.

The first step taken by Russo et al. was to investigate the sensitivity of the results to different model configurations. For this purpose, 30 experiments were performed for PI and MH with different parameter values and perturbed model physics (so 60 in total). The results of these experiments were rather similar, suggesting limited sensitivity. Therefore, it is unlikely that the noted model-data mismatch is related to the set of parameters used. As a next step, short experiments were performed for the MH with prescribed different soil moisture contents in spring. As expected, more humid soils produced lower surface temperatures during summer, in closer agreement with proxy-based reconstructions for Southern Europe. The authors suggest that possibly the use of more complex soil schemes in models could result a longer retention of soil moisture during summer and thus lower summer temperatures. A final analysis concerns an evaluation of the model performance in the 60 different experiments, showing that the optimal performance is different for different variables and periods. This highlights that a model setup that is performing well for the present-day, is not necessary producing an optimal performance for the past or the future.

I expect these results to be of interest for paleoclimate modelers, making this potentially a useful contribution to the literature. However, in my opinion, the manuscript requires a substantial revision before it can be accepted for publication. My main concerns are the following:
A more balanced discussion of the MH summer temperatures in Southern Europe is required. The study assumes that the pollen-based reconstructions are correct, but it is important to make clearer to the reader that there are other proxies that agree with what the models show.

- A few additional experiments with a more sophisticated soil scheme should be performed to make the conclusion of Section 3.2 much more convincing.
- Not all figures and tables provide additional value to the text and could thus be omitted, especially Figure 6 and 7, and Table 4.
- A section should be added to part 3 to discuss the results against earlier modelling studies on the MH climate in Europe.

Further details are provided below.

**Specific comments**

P. 1, line 6. I propose already mentioning here what RCM is used in this study.

P. 2, line 1-2. Please explain the different variables (T2, PRE, TCLC)

P. 2, line 21. Please note that only pollen-based reconstructions suggest cooler summers in S Europe, and there is a good explanation why, see for instance Samartin et al. (2017).

P. 2, line 23. Here references to other climate modelling studies should be included. For instance, Strandberg et al. (2014).

P. 2, line 26. “.... overall simulated warmer conditions, that finds no continental analogue in the proxies”. This is incorrect. The cooler summer conditions at MH are primarily suggested by pollen-based reconstructions. Other proxies (e.g., chironomids, glacier records) do provide evidence for warmer summer conditions during the MH (e.g., Samartin et al., 2017). This point should be discussed more clearly here. It is understandable that the extensive temperature reconstruction of Mauri et al. (2015) is used to evaluate climate models, but it is important to realize that there are also proxies that suggest a contrasting result for summer conditions during the MH. So, there is an alternative explanation to the model-data mismatch, i.e., that the models are correct and the reconstructions are wrong.

P. 2, line 29. Please clarify the dipole mentioned here.

P. 3, line 9. Please provide more information on the overestimation of the summer temperatures. Overestimated by how much?

P. 3, line 14. Is this hypothesis really tested in this manuscript? I would rather say “evaluated”.

P. 3, line 22. “However, this is just an assumption, since there is no guarantee on whether the best model configuration for the present will be the same for other periods of time characterized by different forcing.” In addition, different parameter sets can produce similar present-day mean climates in agreement with observations, but with different sensitivities to radiative forcing perturbations. See for instance Loutre et al. (2011, Clim. Past 7, 511-526). I propose to mention this here.

P. 3, line 24. Please explain RCM and GCM.
P. 4, line 2. Please clarify what “two methods” you mean.

P. 4, line 6: I suggest providing references and an explanation of the COSMO-CLM acronym where it is first mentioned, so here, instead of in Section 2.1.

P. 4, line 22. I suggest mentioning here the calibration approaches to be assessed.

P. 4, line 9. The sentence starting with “Acknowledging the findings...” complex and hard to read. Consider revising.

P. 5, line 3. Please note that although obliquity is an important astronomical parameter, it does not describe changes in the Earth’s orbit around the sun as suggested in the manuscript.

P. 5, Section 2.2. I propose to mention information (including the resolution) on the ocean model here as well. In addition, I suggest mentioning here that the same values for astronomical parameters and greenhouse levels are used in the driving GCM and COSMO-CLM.

P. 5, line 32. On what is “the reference configuration” based? Please explain.

P. 5, line 31. The text mentions 31 experiments, but Table 3 shows 30 different experiments. Is the 31st experiment the reference run?

P. 6, line 19. I do not really see the rationale behind the experiments described in Sections 2.4 and 3.2. As discussed on page 9 (line 10), it is already known for a few years that, when using the default set up, COSMO-CLM has problems retaining the spring soil moisture and that this results in dry soil conditions in summer and anomalously high surface temperatures. So, there is really no need to show that again here. Davin et al. (2016) solved this by applying a more sophisticated soil scheme that performed much better. So instead of the experiments discussed in Sections 2.4 and 3.2 I would suggest evaluating the more complex soil scheme of Davin et al. (2016) for the MH climate. I therefore propose that the authors perform additional experiments with MH conditions and the soil scheme of Davin et al. (2016) and to show the results in figures that replace current Figure 6.

P. 6, line 28. Please provide a bit more information on TERRA_LM. What are the 8 soil types for example, and how do they differ? How is the moisture holding capacity of these soils established? Are the soil types and their characteristics fixed during this study?

P. 6, line 30. The experiments are initialized with 50% relative soil moisture, and then the soil moisture is decreased or increased by 25, 50 or 75%. To avoid confusion, please explain what these percentages mean exactly. For instance: obviously, the initial 50% soil moisture content cannot be reduced by more than 50% if the percentages refer to the same reference soil moisture content. Presumably the initial 50% refers to the volume dictated by the soil water holding capacity, and the 75% refers to “75% of this 50%”, but this is not clearly explained.

P. 7, line 16. Why not assigning all points in the domain? Why leaving some points out? Please explain.

P. 8, Figures 3 and 4. I suggest using a different color scheme for Figure 3. Light red for cooling and green for warming is rather unconventional and could be confusing to the readers. In addition, I propose making the difference between Figures 3 and 4 a bit clearer. As I understand it, Figure 3 shows the mean of the anomalies of MH minus PI for experiments with the same parameters sets. Figure 4 then shows the spread around the
means shown in Figure 3. Are the spreads more or less normally distributed around the mean?

P. 9, Figure 7. I wonder what the added value of Figure 7 is. I can see that the simulation with the best performance (nr 2) for PI has a slightly larger deviation from the reference run in MH, and for MH another simulation than nr 2 is closest to the reference run (i.e., simulation 26). In my view, this can be described in the text without showing the figure.

Section 3.2. In this paper, it is suggested that climate models may simulate too warm summer conditions during the MH because of inadequate representations of soil processes and the inability of models to retain soil moisture during summer. Indeed, Figure 5 shows that summer temperatures are reduced when additional soil moisture is artificially added in MH simulations, which would produce a better match with pollen-based reconstructions in Southern Europe. However, what is not discussed here, is that, at the same time, more humid soils decrease the modeling performance in the northern part of Europe, since here the “standard” result with warmer summers during the MH is in good agreement with proxy-based reconstructions. So, with an increase in soil moisture, summer temperatures would also be reduced here, which is definitively not improving the match with proxies. Consequently, in Northern Europe enhanced soil humidity would not provide a solution. This point should be discussed in Section 3.

Section 3. Discussion: I suggest comparing the obtained results here with reports from earlier modelling studies on the MH climate in Europe and to include a discussion of the impact of lateral boundary conditions on the results presented in this study.

P. 10, Table 4. I suggest omitting Table 4 and simply to describe the information in the text, since the numbers in this table provide little additional value, while the table takes up a lot of space. Besides, the numbers in the table just show the ranking, which does not necessarily provide information on the MAE value of MH versus PI. For example, for T2M, experiment 14 has 2nd rank for MH, but first rank for PI. But this doesn’t necessarily mean that the MAE for MH is lower than the MAE for PI, it just means that there is another experiment that has a higher rank for MH (i.e., 29).

Technical corrections

P. 2, line 22. “proxy-reconstructions” should be “proxy-based reconstructions“.

P. 4, line 8. Should be “feedbacks” instead of singular feedback.

P. 4, line 15. I suggest using “investigating” instead of “supporting”.

P. 5, line 18: one bracket “)“ too many

P. 5, line 20 should be “is applied“

P. 6, line 2: should be “A set of parameters is“ and “affects“ on the next line

P. 6, line 12: should be “from one of the reference simulations

P. 6, line 13: should be “is reported“ (refers to “a list“)

P. 7, line 22. Should be “in both formulas“.

P. 14, lines 29 and 35. The DOI’s of Jungclaus et al. 2012a, 2012b, 2013 are provided two
Figure 2. The colours of number 7 and 8 are not easy to distinguish. I suggest to adjust these colours.

Figure 4. Caption: should be “mean”