

Summary of the manuscript:

Buchan et al. use the PlioMIP2 ensemble to study how the North Atlantic Oscillation (NAO) differs between the preindustrial and the Late Pliocene (MIS KM5c). A factorization analysis on a subset of models separates the contributions of different drivers (CO₂ and geography), and the authors quantify the impact on major weather patterns in Europe. One aspect of the study is to evaluate whether the Pliocene is a meaningful analog for future NAO changes.

General assessment:

I find the research relevant, both for paleoclimate research and for interpreting paleoclimate in the larger picture of climate change. The presentation, however, can still be improved. The figures are of high quality throughout the manuscript and supplement, which is very laudable. My main concern is the organization of the information and the text flow: over large parts of the manuscript, the red thread is missing. On the other hand, I also think that the authors are in parts underselling their work. It could be highlighted more clearly that the reduced ensemble is quite valid towards providing results that are representative for the wider PlioMIP2 framework. See my comments and suggestions below.

=====

MAJOR COMMENTS

=====

M1. Structure and red thread

1. Introduction. Most of the information is relevant, but it should be brought into a more logical order, and three separate sections may not be needed. Please order the information to show (1) what we already know, (2) what we still would like to learn, (3) where current and previous problems lie, and (4) what this study contributes to move us further.

2. Sect. 3. Results and interpretation are tightly interwoven. This may be intentional given the heading of Sect. 3, but I found it difficult to separate the findings of this study from their interpretation and from the link to previous work. There are enough results to justify presenting your own findings and their interpretation in separate sections. I suggest revising Sect. 3 accordingly.

3. Discussion. A clear discussion section is missing (see above).

4. Abstract. Please embed the utility of the Pliocene NAO for future warming better into the reading flow, e.g. by presenting the study as a critical evaluation of the degree of analogy.

M2. Methods

5. Orography and pressure. Eoi400 differs in orography from the modern state. Depending on the model and on how the modelling groups processed their pressure output, this may cause offsets between E280 and Eoi400. I do not expect this to affect the NAO results, but some pressure anomaly plots may differ between models not only because of model dynamics, but also because of different treatments of model output. Please verify that each model's area-weighted global mean pressure is the same in its respective E280 and Eoi400 simulations. If not, I think a standard approach is to correct the Eoi400 mean to be the same as the E280 mean.

6. Lines 135ff (index vs. EOF). Considering EOFs in addition to the index analysis is a good idea. Please state more clearly why both routes are followed and which one is better suited to your work and to the Pliocene.

7. Lines 217ff (internal variability). The NAO has a variability in the order of ~5 years, and the PlioMIP2 modelling groups should mostly have submitted at least 100 model years of data, so internal variability can be tested. This is the standard for equilibrium runs in PlioMIP. If, instead, you want to address with your statement the topics transient forcing, state dependence (e.g. ENSO or AMO), stationarity, autocorrelation, or drift, then please make this explicit. These aspects could also be discussed more explicitly in the revised manuscript.

M3. Proxy and vegetation discussion (Sect. 3.5)

8. Tone of the proxy-data discussion. The discussion reads rather pessimistic and leaves open whether any useful path exists. One could rather highlight here what is already possible and where to go in the future - unless you are really that pessimistic, but then I do not understand the various suggestions that you present where you seem confident that model-data comparison is generally useful.

9. Fixed vs. dynamic vegetation. I think that this aspect of the discussion could be extended. Since you highlight that vegetation has some impact, I as a reader would be interested in your opinion on the use of fixed versus dynamic vegetation simulations and on how each furthers our understanding in different ways.

10. Seasonality and proxy resolution. Please discuss whether high temporal resolution proxy data is really necessary to evaluate model performance on seasonal time scales. Indirect approaches may exist, e.g. deriving from plant types whether springs were wet and summers dry rather than the reverse. A more detailed discussion would increase the impact of your study.

11. Verifiability. Please check that statements are verifiable. The last sentence on p. 16 is supported neither by a reference nor by numbers or findings from your own study.

12. Structure of Sect. 3.5. Please make clear that the vegetation discussion ends on p. 16 and that other proxies follow. This only becomes apparent later and is confusing.

13. Line 294. This statement is not specific to paleoclimate, since it also holds for the present climate. I suggest stating that paleoclimate provides model-independent verification data from periods for which models have not been regularly tuned. Models are commonly tuned to modern climate, which makes verification there difficult.

=====

MINOR COMMENTS

=====

m1. Terminology and consistency

14. Model names. Please use model names consistently across manuscript and supplement, e.g. CCSM4-UoT vs. CCSM4-UofT.

15. Eoi400/E280 vs. Late Pliocene/preindustrial. In most of the main text you refer to Eoi400 as Late Pliocene, but not always. Do you intend Eoi400 and E280 as terms for the simulations, and "Late Pliocene" and "preindustrial" for the periods more generally? If so, please apply this distinction consistently.

16. Fig. 9 / "NAO index". The caption says the NAO index is annotated, but pressure values are shown next to the model names. Strictly speaking, this is the pressure gradient, not the NAO index. Please check the whole manuscript for consistent use of this terminology.

17. Line 30. Should sea ice be listed within a group of large-scale circulation patterns, or should the two be separated more clearly?

18. Line 345. The summer zonal wind speeds are described as "more equatorward". I think you mean that the latitude of the jet (or of the maximum) is more equatorward, which corresponds to a negative Δ latitude.

m2. Methods and framing details

19. Excluded models. Not all models are shown in all figures. Please state for the various figures the reason for exclusion (I assume data availability).

20. Line 86. I found it difficult to locate, in previous publications, how the Rocky Mountains differ between PlioMIP2 and PlioMIP1. Since the authors emphasize their potential impact, please indicate clearly where this information can be found.

21. Line 110. The meaning of "initial assessment" is unclear to me. The PlioMIP2 model output is already well established, so do you plan more detailed work on precipitation shifts in the future? Please clarify.

22. Line 145. Consider adding "despite coarse model grids" at the end of the sentence to highlight the utility of your approach.

m3. Results and figures

23. Fig. 4 (highlighting the validity of your reduced ensemble). For some analyses you use a reduced ensemble because data from other PlioMIP2 models are missing. Nevertheless, Fig. 4 shows in my opinion that your reduced ensemble actually brackets the mean Eoi400-E280 NAO anomaly of the full PlioMIP2 ensemble quite well, best for DJF but also for JJA. For the standard deviation, the coverage of the extremes is weaker but still acceptable. I suggest highlighting this as a strength, which would also preclude potential doubts about robustness regarding the reduced ensemble size.

24. Lines 210 and 216. Please emphasize once more that these statements are based on your results rather than on a different publication.

25. Line 225. The question of variability changes and their impact on European climate should be better linked and discussed. Here it stands somewhat disconnected.

26. Line 266. Does "still" refer to the Pliocene here? Please clarify.

27. Fig. 9. Please consider also identifying the wettest month, to see whether the models diverge strongly. Shifts in precipitation seasonality may be as relevant as changes in its amount. You could refer to Fig. S14 here or include directly in this figure information on the wettest month if meaningful.

28. Fig. 1. Please state whether panels d and h show the same region as the spatial plots or a wider zonal mean. If possible, also show the NAO index boxes in the other maps; this would make the results description in lines 170ff easier to follow.

m4. References

29. Line 18. Please check whether citing Haywood et al. (2024) in the PlioMIP2 context is intended.

30. Lines 23-24. These statements should be supported by a reference. Otto-Bliesner et al. (2017) comes to mind. Recent vegetation modelling work may also be useful, or you could refer to the boundary conditions published within PlioMIP2/PRISM4.

31. Line 42ff. A reference seems to be missing.

==== TECHNICAL COMMENTS =====

General

- Commas. Additional commas would clarify the text flow, in particular after introductory phrases and in complex sentences. Examples: line 290 ("Although negative on average, there are no ..."), line 108 ("To achieve this, data ..."), line 129 ("of the sample period, meaning"). Please also check for misplaced commas, e.g. line 316: "In winter months, on average the, NAO".
- Repetition. Please remove repeated words in close proximity, e.g. "increased" in lines 4 and 5.
- Supplement references. Please ensure that Figs. S6, S7, S9, and S14 are referenced in the manuscript.

By line

- Line 12: fix the grammar ("... but that these differences are may not be related to ...").
- Line 19: fix "from of".
- Line 23: typo "Baring".
- Line 25: fix "above than".
- Line 55: should this be "predicting the future behaviour"?
- Line 59: "it is known that" before "the wintertime" would clarify the formulation.
- Lines 61/62: the wide bracket can be removed; the text reads well without it.
- Line 63 (also line 125): "Models".

- Line 66: fix the circular formulation.
- Line 142: "the the".
- Line 167: the reference should be to S6 and S7.
- Line 176: add a reference to subfigures d and h.
- Line 231: add the closing bracket.
- Line 236: "display" in plural.
- Line 277: "Pliocene NAO on European climate,".
- Lines 284/285: "In northern Europe ... in northern Europe" is repetitive.
- Line 309: fix the grammar. I am also unsure what information shall be conveyed here.

Figures

- Fig. 6: the subfigure references in the caption are misleading. It should maybe read: (a-c) pre-industrial simulation (E280), (d-f) Late Pliocene simulation (Eoi400), and (g-i) the difference between them (Eoi400-E280). Note the missing closing bracket. The same applies to Fig. 7.
- Fig. S2: it seems compressed. Using the full text width would give the y-axes much more space.
- Fig. S14: use "crosses" instead of "crossed", and consider referring to the domain definition for Northern and Southern Europe.

Reference

Otto-Bliesner, B. L., A. Jahn, R. Feng, E. C. Brady, A. Hu, and M. Löffverström (2017), Amplified North Atlantic warming in the late Pliocene by changes in Arctic gateways, *Geophys. Res. Lett.*, 44, 957-964, doi:10.1002/2016GL071805.