CP-2022-59 Davis et al. The climate and vegetation of Europe, North Africa and the Middle East during the Last Glacial Maximum (21,000 years BP) based on pollen data

Dear Editor, Reviewer #1 appears to ask for no further revisions to the manuscript. Reviewer #2 makes 6 comments, which I respond to below in RED:

Referee 2

Thank you for the opportunity to review again the manuscript entitled "The climate and vegetation of Europe, North Africa and the Middle East during the Last Glacial Maximum (21,000 years BP) based on pollen data" by Davis B. and coauthors.

I'd like to thank the authors for taking into account some of my comments in the new manuscript: they've tested the reliability of their approach on the steppe biome and added a table with the R2s and RMSE, and the section on CO2 in the discussion has been greatly improved. Even if I don't necessarily agree with the authors on certain points (I really prefer multi-methods, which is better to understand the reliability of the results obtained), I accept their response.

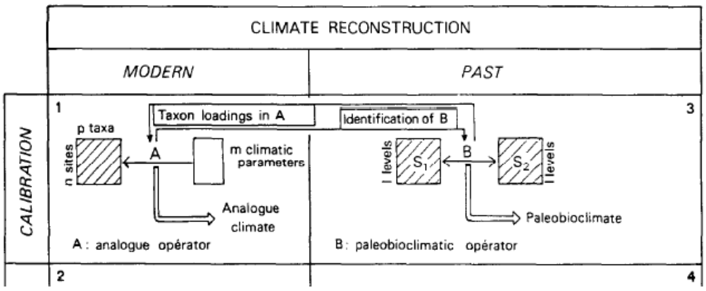
I still have a few comments to make, and as soon as these are taken into account in the next version, I think the paper can be definitively accepted.

-line 220 "to match fossil samples with modern calibration pollen samples": the MAT is an assemblage approach which require no statistical calibration, so correct it (the modern pollen samples dataset is not a calibration dataset as it's the case for the WAPLS for example).
Response: The term calibration is widely used with respect to MAT in the literature. See Simpson (2007) *"The modern analogue technique, described below, is an inverse multivariate **calibration** approach."* Or Juggins & Birks (2012) for instance figure 14.3, part of which is shown below.

1. I don't agree with the authors. The MAT is not a calibration approach : methods based on NN and WAPLS are true transfer function and are based on mathematical calibrations, but MAT is based on a comparison between modern and fossil pollen assemblages (or PFT in your case); there is no calibration in this method (see Guiot et al original paper) and the recent paper by Chevalier et al 2020 "MAT is a classification method, classification techniques compare fossil pollen assemblages to collections of assemblages for which climate is known to identify which assemblages are most similar to the fossil ones".
Please remove the term calibration in the text

Response: The referee is making an argument about semantics. They may be correct in a purely mathematical sense, but words (including scientific terminology) can and do have different meanings in different contexts. The referee may 'disagree' with the use of the word 'calibration' by Simpson (2007) and Juggins & Birks (2012) but the term 'calibration' is widely used by the scientific community when talking about MAT, and especially when talking about the 'calibration dataset' and not 'calibration' as a process. Even in Chavalier et al 2020 in the specific section (5.3.1) on MAT to which the referee refers, the word calibration is used at least twice "..*selecting more and more analogues will progressively include drier samples from the rest of the climate space represented in the **calibration** data, thus inducing an undesired dry bias on the reconstruction (Gajewski, 2015; Viau et al., 2008).* As well as; "*However, including more analogues also increases the risk of false positive matches, especially when the **calibration** dataset encompasses wide spatial areas where the low taxonomic resolution of pollen data can..".*

I am not sure which paper by Guiot "Guiot et al original paper" that the referee is talking about. As with previous comments by the referee, it would be helpful to provide precise information. Certainly, all of the original papers by Joel Guiot from the 1980's that reference an 'analogue' method make reference to calibration, for instance Guiot 1987:



Although the method used by Guiot at this time is not the 'analogue' method that we use, which is closer to Overpeck et al 1985 (in which the term calibration is used at least 4 times).

Action: None

-*-line 221-223 "This is a similar approach to that used by Peyron et al. (1998) and Jost et al. (2005) who also applied pollen PFT scores to reconstruct LGM climate from pollen data, but who used a neural network technique which is a variant of the standard MAT (Chevalier et al., 2020)". I disagree with that, there is a confusion here in the principle of each method. The Artificial neu- ral networks used by peyron et al and others studies IS NOT a variant of the MAT. It's a method close to machine-learning methods, with a real calibration dataset and not easy to check because similar to a black box; in contrast the MAT is very simple, based on an dissimilarity calculation. The only common point is that both methods use PFTs scores to overcome problems associated with the lack of modern analogue but that is all.*

Response: Already agreed, see answer to earlier comment 3.

Action: See answer to comment 3.

2. You have corrected the abstract not this part. Please correct it here too. I propose to replace our sen- tence by "Other methods using PFT scores and artificial neural network techniques have been devel- oped to reconstruct the climate of Europe during the LGM from pollen data (Peyron et al. (1998) and Jost et al (2005).

Response: I am not sure what the reviewer is referring to. The text and line numbers shown above are from the first draft of the manuscript, not the revised version. This section WAS changed in the latest version 3 of the manuscript. These changes were also shown in the response to the reviewer's comments. The text has already been corrected according to the referee's instructions.

Action: None

-line 312 "Similarly, quantitative climate methods have been applied to individual marine pollen records (Combourieu Nebout et al., 2009; Fletcher et al., 2010)": some key references are miss- ing, as the MF Sanchez Goni team. **Response:** Unfortunately, the reviewer does not provide any details of the key references that are supposed to be missing. While MF Sanchez Goni and her team have published many im- portant papers, we cannot find any that involve quantitative reconstructions of climate based on pollen, which is the subject of the sentence

Action: None

3. Salonen, J. & Sanchez Goñi, Maria & Renssen, Hans & Plikk, Anna. (2021). Contrasting northern and southern European winter climate trends during the Last Interglacial. Geology. 49. 10.1130/G49007.1. Or

Sánchez Goñi, M.F., Loutre, M.F., Crucifix, M., Peyron, O., Santos, L., Duprat, J., Malaizé, B., Turon, J.-L., and Peypouquet, J.-P., 2005, Increasing vegetation and climate gradient in western Europe over the Last Glacial inception (122–110 ka): Data–model comparison: Earth and Plan etary Science Letters, v. 231, p. 111–130, https://doi.org/10.1016/j.epsl.2004.12.010.

Response: Ok

Action: The 2 references have been added

- *-lines 337-347 “we did not adjust the pollen assemblage for the over-representation of Pinus in the marine pollen samples” This poses the problem of Pinus transport over very long distances in open environments as the LGM vegetation; this is particularly true for marine cores but it is also true for some terrestrial sites. So the question of excluding or keeping Pinus needs to be more investigated and tested may be on a site-by-site basis*.

Response: Agreed, but the problem of over (or under) representation due to differential transport is a problem that is intrinsic to the science of palynology with no straight-forward answer. Fundamental to this is the fact that although the risk of under/over representation can be acknowledged, it is generally very difficult to detect and correct in any detail. One of the closest attempts at this can be seen in the use of MAT methods to reconstruct tree cover, which we apply in our study. This ‘black box’ approach at least makes some attempt to take into account the potential over-representation of *Pinus* in both terrestrial and marine environments, at least where this problem is also found in the modern analogue samples that are matched with the fossil samples.

From the point of view of the marine pollen records, we find it more appropriate to include rather than exclude *Pinus* because it is a key forest forming tree in the coastal regions close to the marine sites and to remove it completely would create an artificially arid assemblage that would certainly undermine the ability of the transfer function to reconstruct precipitation. Reconstructions of temperature would be less affected because *Pinus* is a generalist found in both hot and cold regions and so carries only a weak temperature signal compared to the rest of the assemblage. We can add this clarification to the text.

Action: The following text has been added to clarify the problem for the reader (lines 435-439) “***Removing Pinus from the assemblage would almost certainly create an artificially arid assemblage in these circumstances, undermining the ability of the transfer function to reconstruct precipitation, although temperature would likely be less affected since Pinus is a generalist found in both hot and cold temperature regions***.”.

4. I don’t agree with the fact that you keep Pinus in the marine records for the climate reconstruction: it could change the biome! Palynologists working on marine records exclude it of the pollen sum, and its particularly true for open environment as LGM. A solution may be you to test your method with and without pinus and check the incidence of removing Pinus on the climate reconstruction by comparing the results with terrestrial close records. It’s an important point for me, but I leave the editor take his decision.

Response: The referee has not engaged with the argument. Removing *Pinus* can be just as detrimental as keeping *Pinus* (and it can equally “change the biome”), they are two sides of the same problem. In the manuscript we give the example of coastal Portugal where *Pinus* is the dominant forest forming tree. If you remove *Pinus* from the sum then you are removing a key component of the ecosystem, and a key indicator of the climate. It should be enough that the reader is reminded of this problem, which we do in the text. I would note that both Salonen et al 2021 and Sánchez Goñi 2005 highlighted by the referee (see above) don’t even mention this problem, and it is not clear that Salonen et al 2021 even removed *Pinus* in their marine-based pollen-climate reconstruction, since the method section simply says that they used the same method as an earlier paper, and that earlier paper only analyzed terrestrial samples where *Pinus* <u>was</u> included in the sum.

Action: None

-lines 443-444 “the neural-network methodology of Peyron et al. (1998) and Jost et al. (2005) which we call MAT-NN, as well as the Inverse Modelling approach by Wu et al. (2007) which we call INV.” First, the neural networks methodology of peyron et al. is NOT a MAT method, so you cannot call it MAT-NN, it’s a non-sense. Second, could you use the name of the method given in the reference papers? Please check, I guess it’s the PFT method for Peyron et al and I.M. for Wu et al. which are correct.

Response: We intended our method acronyms to be as self-explanatory as possible. 'PFT' is not the defining feature of the Peyron et al 1998 method, since the use of PFT scores can, and has, been used in other methods such as MAT. We therefore prefer to use the acronym 'ANN' for Artificial Neural Network (as used by Chevalier et al 2019).
Action: The section mentioned by the reviewer has been moved to the discussion as they suggest. MAT-NN has been changed to ANN throughout.

5. Better to use PFT-ANN for the study of Peyron et al and IM for the study of Wu et al (not INV, I dont understand why you have changed it): please correct in the text

Response: I don't really see the necessity to add PFT to ANN since we only use the ANN results of Peyron et al, who used PFT scores in their ANN analysis. Peyron et al themselves use just the acronym ANN, as do Chevalier et al 2019. Confusingly Wu et al call the ANN approach of Peyron et al just 'PFT'. Wu et al also call their inverse modelling approach IVM, not IM as the referee suggests. In summary there does not appear to be a consensus on the use of acronyms for these methods, although more so for ANN, so the fact that we use ANN and INV seems reasonable.

Action: None

-lines 615-616: "expected, areas of forest reconstruct similar or increased precipitation compared to today, and areas of steppe indicate deceased precipitation (see next section)." The CO2 effect on climate reconstruction (see recent papers by Cleator et al. and Prentice et al) is not discussed, please add a part on this point.
Response: Ok
Action: The CO2 problem is revisited in the discussion (lines 852-869)
Ok, perfect

-lines 778-784: Good to add a comparsion with the brGDDTs temperature record from Padul (Rodrigo-Gámiz et al., 2022).
Response: We are reluctant to include this study by Rodrigo-Gámiz et al. 2022 because this record looks quite odd. In particular, it appears warmer than the present day for much of the glacial period and has a long-term trend very similar to pH. This is important because the brGDDT proxy has been criticised for being influenced by pH as well as temperature, although this potential bias does not appear to be mentioned in the paper. We do not think that excluding the study would make any significant difference to the conclusions of the paper.
Action: None
6. It's often the case with BRGDGTs studies: the temperature values are depending on the calibration used (here the Martinez-Sosa et al one), and are often too high. I think that the most important is to look at the climate patterns: an important result is that they show that LGM temperature were higher than those reconstructed during Heinrich events. Please cite this paper.

Response: The referee is insisting on us citing a paper that has little or no scientific merit in relation to our manuscript. I do not know if the referee has any connection with this paper, but this appears to us to be highly unethical.

Action: None