Response to reviewer

Our responses are given in blue and proposed changes to the text in *blue italics*. References to line numbers refer to the revised manuscript.

1. The last paragraph of the introduction should highlight the two main goals of the papers (methodological test and Holocene climate improvement such as explain in the general observations).

The methodological test was originally presented as a minor aspect of the paper for a good reason, namely that the modification made is relatively slight. It concerns the algorithm used to implement the sampling frequency correction, without changing the underlying principles of the method as presented (and extensively tested) by Liu et al. (2020). The modification does nonetheless produce a further improvement in reconstruction accuracy, which we agree is worth greater emphasis. We have accordingly modified the last paragraph of the Introduction as follows:

We used the method Tolerance-weighted Weighted Average Partial Least-Squares regression with a sampling frequency correction (fxTWA-PLS), introduced by Liu et al. (2020) as an improvement of the widely used Weighted Average Partial Least-Squares (WAPLS: ter Braak and Juggins, 1993) method for reconstructing past climates from pollen assemblages. As presented in depth by Liu et al. (2020), this method is a more complete implementation of the theory underlying WA-PLS because it takes greater account of the climatic information provided by taxa with more limited climatic ranges and also applies the sampling frequency correction to reduce the impact of uneven sampling in the training data set. Liu et al. (2020) showed that fxTWA-PLS does indeed provide better reconstructions than WA-PLS.

Here we have further modified the algorithm implementing fxTWA-PLS, achieving an additional gain in performance. In the algorithm as published by Liu et al. (2020), sampling frequencies were extracted from a histogram. In the modified algorithm they are estimated using P-splines smoothing (Eilers and Marx, 2021), which makes the estimates almost independent on the chosen bin width (see Appendix A for details). In addition, the modified method applies the sampling frequency correction at two separate steps – the estimation of optima and tolerances, and the regression step – a measure intended to produce more stable results. Indeed, the modified method produces both improved R^2 values and reduced compression and maximum bias in reconstructed climate variables (see Table A1 and Figs A1–A2). We will return to this point in the Discussion.

We have used this improved method to reconstruct Holocene climates across Iberia, and re-examined the trends in summer and winter temperature and plant-available moisture, using a new and *relatively comprehensive compilation of pollen data (Shen et al., 2022) with age models based on the latest radiocarbon calibration curve (IntCal20: Reimer et al., 2020)…*

2. The methodological part is far richer and clearer than the previous version of the manuscript with a larger and more exhaustive list of existing methods to convert the pollen signal to climate parameters. However, we do think that some point in the methodological choice should be clarified in the manuscript (maybe in introduction and certainly in discussion) and all these choices have to be defended in introduction / methods:

2.1. How did you select the studied climate parameters? Why MTWA, MTCO and alpha instead of MAAT, MAP, GDD_0, etc? This is an important point.

We have modified the Methods as follows:

There are no generally accepted rules as to the choice of variables for palaeoclimate reconstruction. No systematic comparison of these choices has been made. However, it is widely understood that plant taxon distributions reflect distinct, largely independent controls by summer temperatures, winter temperatures, and moisture availability (see e.g. Harrison et al., 2010). Therefore, in common with many other studies (Cheddadi et al., 1997; Jiang et al., 2010; Peyron et al., 1998; Wei et al., 2021; Zhang et al., 2007), we have chosen bioclimatic variables that reflect these independent controls, with mean temperature of the coldest month (MTCO) to represent winter temperatures, mean temperature of the warmest month (MTWA) to represent summer temperatures and α, an estimate of the ratio of actual evapotranspiration to equilibrium evapotranspiration, to represent plant-available moisture. We choose not to use mean annual air temperature (MAAT) because it is a composite of summer and winter conditions; and we prefer to use an index of effective moisture availability (our estimate of α being one such index) to mean annual precipitation (MAP), whose significance for plant function depends strongly on potential evaporation (a function of temperature and net radiation). Our calculation of α takes account of this dependence. Growing degree days above a baseline of 0 °C (GDD_0) would be a possible alternative to MTWA as an expression of summer conditions but is most relevant as a predictor of "cold limits" of trees in cool climates, whereas MTWA better reflects the high-temperature stress on plants in Mediterranean-type climates.

2.2. About the independence between climate parameters, it is also not really clear. The CCA and VIF show than the climate parameters are independent but in the same time you show Fig. 6 and 1. 212-214 than they are closely correlated. This have also to be discussed.

The CCA and the VIFs do indeed show that there is sufficient independent information in pollen data to allow all three variables to be reconstructed. In other words, if their true values varied independently in the past, we would expect these variations to be manifested in the fossil pollen data – and we would expect the reconstruction to reveal them. VIFs are the standard metric to assess such independence. Acceptable VIFs can be achieved even if there is some correlation between variables in the training set (although very high correlations would be problematic, leading to high VIFs).

Fig. 6 shows a quite different point. It shows that there was, in fact, a high correlation between the variations of α and MTWA in this region during the period studied (panel b, which reflects both spatial and temporal correlation)– which we can then interpret in terms of potential climatic mechanisms. Panel a only reflects spatial correlation and the correlation is acceptable to reflect independent information, showing that the relationship in past reconstructed values of the two variables is not an artefact of correlations existing in the training data set.

We have modified text as follows:

The variance inflation factor (VIF) scores are all less than 6, so there are no multicollinearity problems (Table 3) (Allison, 1994), making it possible to independently reconstruct all three climate variables based on pollen data.

2.3. Why using 10° resolution climate database instead of already interpolated and discuss climate databases with 1° resolution (such as WorldClim2 or CHELSA data based)?

We obtained the climates for each site using geographically weighted regression (GWR), allowing a built-in correction for site elevation. By far the most important difference between CRU and the more recent, higher-resolution data sets is simply that they account for elevation with much finer resolution. However, even 1 km resolution is not necessarily sufficient to pinpoint the climate of a site; GWR would still be needed. Moreover, other work in Sandy Harrison's group has confirmed that so long as GWR is used, the use of higher-resolution data results in only minimal changes to the interpolated site climates.

We have modified text as follows:

The climate at each site was obtained using geographically weighted regression (GWR) of the CRU CL v2.0 gridded dataset of modern (1961-1990) surface climate at 10 arc minute resolution (New et al., 2002) in order to (a) correct for elevation differences between each pollen site and the *corresponding grid cell and (b) make the resulting climate independent of the resolution of the underlying data set.*

3. About the results and discussion also some modifications are necessary. We think that the text, especially in the discussion is not sufficiently connecting with figures. Figs. 1, 5 and 6 are called only once and all the Figs. 1 to 6 are only called in results and not in discussions.

We intended to separate the result and discussions, now we have included more reference to them at line 278, 279, 305, 343, 344, 349, 353, 354 at the revised manuscript without tracked changes.

4. About the discussion, only the last paragraph of the discussion focuses on the improvement made with the fxTWA-PLS2 version of the transfer function. The first sentence of the conclusion is "We have developed an improved version of fxTWA-PLS which further reduces compression bias and provides robust climate reconstructions", however this as not be proved neither discuss in the manuscript. We argue that this topic should be discussed and validated in the first part of the discussion before presenting the climate composite reconstruction and comparing it with other proxies, climate modeling and so on. Especially using the material in Appendix (Table A1 and Figs. A1 and A2).

We have now modified the beginning of the Discussion as follows, in response to this request.

The modified version of fxTWA-PLS (fxTWA-PLS2) (Table 2, Table A1) shows a few differences compared to the previous version (fxTWA-PLS1). Cross-validation R^2 values are higher for MTCO and MTWA, and almost unchanged for α. The maximum bias shows a decrease for all the three variables, especially for MTCO. The compression problem is also reduced for MTCO (b_1 increases from 0.82 to 0.91) and MTWA (b_1 increases from 0.69 to 0.71) while remaining roughly the same for α. The overall performance statistics thus show substantial improvements for MTCO and MTWA, while they show little change for α. However, Figure A1 shows that "unphysical" reconstructions beyond the natural limits of α (0–1.26) are greatly reduced, especially for the lower limit. There are also fewer outliers in Figure A1 and A2 for all three variables. Thus overall, the modified version further reduces the reconstruction biases, especially at the extremes of the sampled climate range. This improvement probably occurs because of the separate application of 1/fx correction during both the calculation of optima and tolerances of taxa and during the regression step – instead of applying an overall weight of $1/fx^2$ at the regression step, which can result in some extreme values (with low sampling frequency) being weighed too strongly and appearing as outliers.

fxTWA-PLS2 reconstructed climates have shown that there was a gradual increase in …