

We note that the reviewer has uploaded the same comments twice, although the formats are changed. Since we assume this was an error, we only provide a single response here.

Original comments are in **bold**, our response in *italics*, with suggested changes in the text in normal script.

**Line 41: “long-term” rather than ‘longterm’**

*We will change this.*

**Line 49: “ice-sheet volume” rather than ‘ice sheet volume’**

*We will change this*

**Line 66: Valero-Garcés et al. needs a year**

*We have added 2019. This paper is accepted for publication.*

**Line 63: Two of these papers (González-Sampériz et al. 2013; Aranbarri et al. 2014) present parts of the entire pollen sequence from Villarquimado (V). It would be a great help to the readers of this manuscript to have a more detailed complete pollen diagram than Figure 2 that only shows four groups of taxa.**

*It is true that the Aranbarri et al. (2014) only covers the last 13,000 years, but González-Sampériz et al. 2013 provide a simplified diagram of part of the whole sequence. A more detailed paper describing the pollen sequence and vegetation history from Villarquemado is currently in preparation. The goal of the current paper is to provide a climate reconstruction rather than to document the pollen record. We therefore showed Poaceae and Polypodiales in order to illustrate the spiky nature of the record –which affects our reconstructions. Similarly, we showed the biodiversity index because we use this to explain why some samples do not provide good reconstructions. The major changes of climate are reflected in the changes between Mediterranean and Steppe plant functional type, which is why we show these. However, we will expand Figure 2 to include more key taxa groups and provide more comprehensive information regarding the palynological sequence.*

**Line 66: Again it would be useful to see the age depth model for V so that one can judge ‘the quality of the age model’.**

*The Valero-Garcés et al., which describes the construction of the age-depth model in great detail is accepted for publication and is expected to be available within the next few weeks. However, we will include the Bayesian age model for the sequence, derived from this paper, in our SI.*

**Lines 102-113: I assume that all the data not already in public databases such as EMPD, EPD, and Pangaea will soon be publicly available (e.g. some of the EMBSecBIO data that are not in the EPD).**

*The modern surface pollen data set that we use for these reconstructions is publically available and we will give the DOI for the data set in the text. The fossil data will be made available via Neotoma and/or EPD upon publication of the paper currently in preparation describing the palynological sequence in detail.*

**Line 139: Were these GAMs constructed using pollen percentages? If so, what**

**taxa were included in the calculation sum?**

*We will rewrite this text as follows:*

The GAMs were implemented with the mgcv R package (Wood, 2017). The R implementation makes the selection of the smoothing parameters automatic (Guisan et al., 2002). We used pollen percentages based on a pollen sum that included all of the 196 taxa.

**Line 148: It would have been useful to present 2–4 of these GAMs.**

*Since these are large figures, we will include two examples of the GAMs in the Supplementary Material (Supplementary Figure 1 and 2). We will also change the figure numbering in the text.*

**Line 156: Were spores of pteridophytes and lycophytes included in the pollen sum?**

*The pollen sum includes all of the taxa that were used for the reconstruction, please see Supplementary Table 2.*

**Line 218: Citation needed for the statement about WA-PLs being ‘relatively robust to spatial autocorrelation’ (e.g. Telford and Birks 2005 or 2008 QSR).**

*We will add a reference here.*

**Line 226: ter Braak et al. 1993 is missing in the reference list**

*Thanks for pointing this out. We will add it into the reference list.*

**Lines 221-227: Why use the Van der Voet t-test to select the number of components and then use the lowest RMSEP as a basis for selecting WA-PLS components? The Van der Voet test is less prone to lead to over-fitting (see Juggins and Birks 2012).**

*The selected component was first determined on the basis that (a) the Van der Voet test showed it was significant, whereas subsequent components were non-significant and then (b) by choosing the lowest RMSEP of the significant components. We obviously did not explain this clearly enough and will rewrite the text as follows:*

The performance of the calibration models was assessed through leave one-out cross validation. The number of components used in each model was estimated through a randomisation t-test on the results of this cross validation (Van der Voet, 1994). We selected the significant component with the lowest root mean square error of prediction (RMSEP), but only if there was a significant improvement in RMSEP relative to a lower number of components – since including more components can result in over-fitting of the data so that model predictive value decreases (ter Braak et al., 1993). We checked that the final transfer functions had a high R<sup>2</sup> for prediction and a low maximum bias.

**Line 239: Is there a word missing after ‘effects’?**

*This should have read “effects of” and we will correct this*

**Lines 258-261: The CCA (Table 1 lower half) does not show that the three climatic variables have an independent contribution to explaining variation in pollen abundances as you do not appear to have conditional (partial) CCAs with MI as the predictor variable and MTCO and GDD0 as covariables. Also it should be “Axes” not ‘Aexs’.**

*We have corrected the typo.*

*We have run a partial CCA as suggested, with each climate variable as the predictor and the others as covariables. This shows that they all have an independent and significant ( $p < 0.001$ ) contribution to explaining the variability. We will add this information in the paper.*

**Lines 263-264: Three and four components are quite high – does the Van der Voet test show that these components are significant?**

*As the P values in Table 2 show, these components are all significant at the 95% confidence level. We have also checked the impact of using component 3 for MTCO and GDD, and component 2 for MI (i.e. the last components that are significant at 99% confidence rather than 95% confidence), and found that this does not affect the reconstructions noticeably or change the conclusions of the paper. We will include these alternative reconstructions in the SI.*

**Line 269: But Polypodiales are not really aquatic. Of the taxa you list in S1 Table 2, Thelypteris palustris is the only mire plant. None are aquatics.**

*We have expressed this badly. We think that the distinct Polypodiales peak values are the result of inwash into this semi-aquatic environment (as shown by the nature of the sedimentary record) because the spores are persistent in the environment. We will rephrase this as follows:*

*These correspond to samples that have unusually high values of either Poaceae or Polypodiales (Fig. 2), and where the sedimentary record indicates that environmental conditions were fluctuating and the basin was occupied by wetlands. Under these conditions, the anomalously high peaks of Poaceae likely correspond to reeds and the anomalously high peaks of Polypodiales could represent inwash.*

**Line 480: Is there a word missing after ‘regarding’?**

*This should have read “regarding the” and we will correct this.*

**Lines 542-554: A corrected list of authors of this paper was published as a correction in VHA.**

*We will indicate that there has been a correction to the paper in the reference list.*

**Line 623: “Quaternary” not ‘quaternary’**

*We will correct this*

**Lines 823-828: Maps a and b are very small and b is impossible to read.**

*We agree that these two maps are very small. Figure a is not necessary for this paper and we will remove it and we will redraw Figure c to make the location of Villarquemado clearer. We think that Figure b is helpful in understanding the setting of the site. We will therefore preserve this figure but treat it as a separate figure from plot c. We will renumber the figures in the main text accordingly.*

**Figure 4: It is a missed opportunity not to have used bootstrapping (available in rioja) to estimate sample-specific errors for these three reconstructions.**

*The bootstrapped errors as calculated in rioja are large (MTCO = 4.8°C, GDD = 900, MI = 0.4) but these values are comparable to other WA-PLS reconstructions (e.g. Lake Ohrid, MAT = 5°C, PANN = 200 mm, see Sinapoli et al., 2019). The issue here is that the standard leave-one-out bootstrapping approach provides a measure of the*

*stability of the model to the individual samples in the training data set. We have applied an alternative approach of creating 1000 training data sets by random selection and then examining the stability of the taxon coefficients. We then apply these different models to estimate the reconstruction error associated with individual samples in the fossil record. In this approach, samples which are dominated by taxa with stable coefficients will have relatively narrow confidence intervals while samples dominated by taxa with non-stable coefficients will yield large errors. We believe this approach gives more realistic confidence intervals for individual samples than the standard method. We will update the figures to include these confidence intervals.*

**Table 2: These do not seem to me to be the results of the Van der Voet randomisation t-test. Also here it says that Poaceae and Polypodiales (not correct spelling) were omitted but on line 159 you say 196 taxa were used. What were actually used – 194 or 196 taxa?**

*The P values are indeed from the Van der Voet randomization t-test. We will modify the caption to make this clear as follows:*

The results of the leave-one-out cross-validated predictions of the weighted averaging-partial least squares (WA-PLS) regression models used for the climate reconstructions. The P values are derived from the randomisation t-test on these results. The final model is based on 194 taxa, omitting Poaceae and Polypodiales. Selected components in the final model are marked in bold.

*We originally used 196 taxa to do the reconstruction but, as explained in the results, we found that Poaceae and Polypodiales had highly anomalous peaks that influenced the reconstructions and we therefore left them out in the final reconstruction. As we point out, this influenced the reconstructions for the samples with anomalously high Poaceae and Polypodiales but did not affect the reconstructions for the other samples. Nevertheless, as we present results both the reconstructions with and without Poaceae and Polypodiales (Supplementary Table 3) we need to explain which results are based on 196 taxa and which are based on 194 taxa. We will modify the description in the methods (Line 158-159) to read:*

Amalgamated taxa that occur in less than 10 sites were not considered in the analyses, reducing the number examined from 249 to 196 taxa. We also ran further analyses after removing two taxa that displayed anomalous behavior. The final results are therefore based on 194 taxa.

*We will also modify the text at Line 270-272 to read:*

Both Poaceae and Polypodiales were therefore removed from the final WA-PLS model (Table 2), reducing the total number of taxa considered from 196 to 194.

**SI Table 2: I am not sure that this is needed when much more relevant basic data are not given.**

*As explained above, basic information on the age model and the pollen record are given in other publications. We will include a figure in the SI with the age-depth model and we will expand the presentation of the summary pollen diagram to include more functional types. However, our focus here is on the climate reconstruction, and we think it is useful to have this table to explain how taxa were amalgamated to make the reconstruction.*

*It is true that some pollen types that we have grouped together are morphological distinct. However, our data set was compiled from multiple different published data*

*sets and not every palynologist made these distinctions. Some morphologically distinct pollen types are only recorded at a very small number of sites and often these sites are geographically clustered, suggesting a systematic recording bias. Including taxa that are only recorded rarely (or show geographic clustering suggesting that they have not been sampled across the whole of their potential climate range) would lead to misleading reconstructions. Niche conservatism would suggest that higher taxa should have a coherent environmental distribution (and this is more reasonable than the idea that a morphologically distinctive pollen type would have a coherent distribution). Furthermore, we checked the climate space of component taxa that were amalgamated into higher taxa using the GAMs to ensure that the grouped taxa showed similar climatic preferences.*

- a) Asteroideae have very distinct pollen (Ambrosia, Xanthium). *Yes, but Ambrosia is only identified <3% of the sites and Xanthium only <5% of the sites in the data set. According to the GAMs, they occupy the same climate niche*
- b) Cardioideae, Caryophyllaceae (Spergula, Paronychia, Illecebrum, Agrostemma). *The same is true of these species: Spergula (<2%), Paronychia (<0.1%), Illecebrum (<0.1%) Agrostemma (<0.1%)*
- c) Succisa is not in the Caryophyllaceae. *Succisa is in the Caprifoliaceae and this was a mistake in our Table and we will correct this.*
- d) Fabaceae and Fabaceae (herbs) is an unsatisfactory division. *We agree this is unsatisfactory. However, at a large number of sites herbaceous taxa have been identified separately from Fabaceae as a whole. The identified herbaceous taxa are usually rare and would not be used unless they were grouped together, and we checked they occupied a similar area of climate space and that this was distinctive from the climate space occupied by the taxon that was simply called Fabaceae in the same samples.*
- e) Tuberaria pollen is like Helianthemum. *Helianthemum is identified very commonly in the data set (12% of the sites), whereas Tuberaria is only found at 0.03% of the sites. We could also have included Helianthemum in the Cistaceae, but this seemed like a loss of information given that many palynologists have identified it explicitly.*
- f) Hepatica and Pulsatilla pollen are like Anemone pollen. *It is true that these are rare and we have amalgamated them into Ranunculaceae and re-run the reconstructions. The change has no effect on the results.*
- g) Hypericaceae pollen are the same as Guttiferae = Clusiaceae. *Hypericaceae is an accepted family and Hypericum is explicitly identified at 252 sites in the data base; Clusiaceae (on the other hand) is a separate, recognised family but representatives are extremely rare in the data set and therefore not used in our reconstructions.*
- h) Hippuris (in Plantaginaceae) is an aquatic whose pollen is totally different from Plantago pollen. *We agree that Hippuris, which we have put into the Plantaginaceae, is an aquatic and should be omitted from the data set. It only occurs in a few sites and omitting this does not affect our reconstructions.*
- i) Cryptomeria is a tree, not a fern (Pteridaceae), many of the Pteridaceae have distinct spores (e.g. Adiantum, Cryptogramma). *The inclusion of Cryptomeria into the Pteridaceae was a mistake in the table and we will correct this. It only occurs in one site in the data set, and so will not be used in the reconstructions.*
- j) Rosaceae (Fragaria has Potentilla-type pollen). *Fragaria only occur at 5 sites in the data set but Potentilla occurs 906 times and furthermore occupies a distinctive climate space. We could also have included Potentilla in the Rosaceae, but this seemed like a loss of information given that many palynologists have identified it explicitly.*

k) some of the taxa in your Scrophulariaceae are now in the Orobanchaceae (Euphrasia, Pedicularis, Rhinanthus). *Yes these three genera are now in the Orobanchaceae, as in fact is Melampyrum – which we also included in Scrophulariaceae. We will remove these from Scrophulariaceae and include Orobanchaceae as a new taxon. They only occur at a few sites and this change has no impact on the reconstructions.*