

I am grateful for the opportunity to review the article entitled “Examining bias in pollen-based quantitative climate reconstructions induced by human impact on vegetation” by Wei Ding, Qinghai Xu, Pavel Tarasov, for *Climate of the Past*.

I think this paper is well written, very clear and concise, and sufficiently interdisciplinary in content to be published in *Climate of the Past*. This study focuses on two modern pollen data calibration sets, one considered as “natural”, one with high human impact. These 2 modern pollen datasets are tested to reconstruct the annual precipitation changes during the late glacial and the Holocene in north-central China, at the margin of the East Asian Summer; they used the classical Weighted Averaging Partial Least Squares (WA-PLS) approach to reconstruct quantitatively the precipitation.

This study is very interesting given that the problem of human impact on pollen data is a key problem in climate reconstructions. I think that the paper of Ding et al. presents interesting findings in terms of results; in particular they support the consensus that climate reconstruction based on pollen data must be taken with caution for recent periods (Bronze Age to 0 ka).

So, I recommend the publication with minor changes, which are listed below.

### **Main points**

- I think that test two different datasets -, one considered as “natural”, one with high human impact - is really interesting. But in fact, transfer functions to reconstruct past climate are based on the principle that vegetation and climate are in equilibrium-which is not always true, I know-. We keep only pollen samples which have been collected in « natural » ecosystems, we exclude samples which have been collected in areas with high human impact. One another test interesting to do is to test the reconstruction with an unique dataset (natural and non natural) to test if the results will be different or not because we will never use a dataset with human impact.
- what is the definition of « natural vegetation » given the changes made by human societies on their environment , particularly during the late Holocene? How do you define natural vegetation ? is it potential vegetation?
- your study takes only into account the bias linked to modern pollen datasets ; what about the potential biases on fossil data linked to past human impact (fires ...); could you more discuss this point?
- I have not seen a description of the fossil record used. Even if it has been already published, it's an important point. Could you add a simplified pollen diagram for the fossil pollen sequence? Could you give details on pollen assemblages? Or describe it very briefly? Time period covered?
- Did you calculate error bars? I have not seen them in the figure. Error bars are needed before publication to discuss more in depth the possible bias with the H set. And to see if the differences between the 2 climate reconstructions are significant or not.

### **Minor points**

1. Title: could you add “in China” in the title?
2. The abstract is well written and informative. May be you can precise in the abstract that the bias in the pollen –based reconstruction is important only during recent periods (not entire lateglacial and Holocene)
3. Introduction
  - p.1, line 26: the ref Von Post, (1916) is more appropriate

- p 1, line 30: I don't agree with the sentence: Palaeoclimatology relies on modern pollen-climate relationship studies; Palaeoclimatology is based on pollen but also on various proxies (speleothems, lakes, tree-rings...) not only pollen; please correct.
- p.1, line 32-33: the ref Bartlein et al., 2011 is missing
- p.1, line 36: what is "the principle of uniformitarianism"?
- p.2, line 5: "in such regions": which ones?

#### 4. Material and methods

- p.3, line 28: how many pollen data do you use from each region and ecosystem types?
- p.3, line 37: how do you define the "natural vegetation communities? Do you have pollen traps and vegetation surveys?
- p.4, line 6: is the pollen sum enough high after exclusion of anthropic taxa, for example in cultivated lands? line 7: it will be informative to precise how many pollen spectra you have per ecosystem (both dataset).
- p.4, lines 11-12: do you also have calculated GDD5 and the moisture index (prentice et al...)?
- p.4, line 14: how do you correct the values for the precipitation parameter?
- p.4, lines 18: more details on the calculation of HII are needed: HII is based only on modern human impact, it doesn't take into account the human impact during recent periods (from 2800 to 0).
- p. 4 part 3.3 see main point 4
- p.5, line 3: which statistical techniques have been tested? Could you write 1 sentence on the concept of the WAPLS?
- p.5, line 14: I don't understand, you also use MAT? Why? Please explain more and give more details on the methods

#### 5. Results

- p.5, line 21: "the differences between 151(147) and 99(93) taxa in the Natural (human) set are only explained by the exclusion of rare taxa?"
- p.5, line 38: I don't understand why you don't keep the 3 component model (even given the threshold of 5%).

#### 6. Discussion

- P 5, line 31 ..... is a function of climate : yes but also other factors play a role
- p. 5, line 18: a significant bias: do you mean statistically significant? it has been tested?
- p.6, line 16: how is calculated the " mean analogue-HII values in the N-set"?
- p. 6 line 32: "This raises the question of whether the Holocene climate could be quantitatively reconstructed using pollen data from eastern China. The answer is not a simple 'yes' or 'no'"; I think that the question is "when" and "where" given that your reconstruction before 2.8 ka can be considered as robust and climate driven

#### 7. Tables and figures

- Fig 1: a, b, the legend is not clear (the right one); the site is not easy to find on the fig: please correct