I am grateful for the opportunity to review the article entitled "Quantitative reconstruction of precipitation changes in the Iberian Peninsula during the Late Pleistocene and the Holocene" by Liisa Ilvonen and co-authors.
This study aims to reconstruct, from pollen data, the precipitation changes in Spain during the LateGlacial and the Holocene with a multi-method approach. Results show different climate patterns during the Holocene.

I think that the paper of Liisa Ilvonen et al. presents interesting new findings in terms of results to be published in Climate of the Past but I also think that it cannot be published in its current version for several reasons.

I first suggest to better highlight the innovative side of this study. In particular if we compare the objective of this work with those of the paper of Tarroso et al (2016) which focus on the reconstruction of the climate (Temperature and precipitation) in Iberian peninsula during the last 15000 years from pollen data. This study has several positive points that should be further highlighted in the text: a new modern pollen dataset, a multi-method approach… However, the description of the modern pollen dataset is too short and the discussion on the multi -method approach needs to be improved. The discussion is essentially based on the results of the WAPLS: why? This point must be justified. If the results of the Bayesien method are not robust, then yes, you can only discuss the WAPLS, otherwise you have to discuss both.

My second point concerns the lack of comparison of your results with the precipitation curves available in the Mediterranean area: the study of Tarroso et al (2016) for Spain, Dormoy et al (2009) for south Spain; studies of Peyron et al. (2011; 2013), Combourieu-Nebout et al., 2013 and Magny et al (2013) for Italy. It's important to add these curves in the figures (6 or 7?) to discuss the regional climate pattern. Particularly the curves of Tarroso et al., (2016) which are based on another climate reconstruction method, the PDF, show clearly a different pattern than the precipitation reconstructed here; the differences have to be discussed more in depth.

Third point: this study does not propose a temperature reconstruction inferred from pollen: why? The authors assume that precipitation is the most important climate parameter, but this is not justified in a statistical point of view. Multivariate analyses would be required to prove the role of annual precipitation. Temperature curves based on the same pollen data should be added, or their absence should be statistically justified; Authors can also plot the temperature curves inferred from pollen data as in Tarroso et al. 2016. Much of the discussion and figures are based on chironomidss temperature curves (figures 6, and7), so either the authors remove the temperature curves to base the discussion only on precipitation (and compate it with more regional precipitation patterns), or the authors apply their methods to produce temperature curves, or the authors include the temperature curves of Tarroso et al for Spain.

Last point: Authors don't investigate the links of these reconstructed climate changes with the different climate forcings. It's an important missing point.

Other points

- Data sources
    - The paragraph on the modern pollen dataset is too short given that the quality and accuracy of the modern pollen dataset is very important in transfer functions. The modern pollen dataset used here has never been published, so more details are needed: could you add a table or a map with the biome corresponding to each modern sample? We need it to be sure that all the vegetation type occurring in the past are included in your dataset. I particularly think about the more herbaceous

during the Younger Dryas and the taxa of the Bolling/Allerod. Another important point to discuss is the human impact: how do you deal with that in the modern dataset? Do you exclude anthropic taxa?
    - The choice of the 7 fossil pollen data should be explained and justified; for example, why did you only choose Quintanar de la Sierra, as the sequence which cover the Lateglacial while other pollen records are available in the EPD?

- Reconstruction of past variables
    - Line 26: you test the performance of the calibration of the transfer function, you don't test the performance of the modern training set: please correct.
    - Line 28: reformulate: for constructing the transfer functions for annual precipitation
    - The WAPLS is a classic method, often used in paleoclimate studies. In contrast, the Bayesien reconstruction method is newer: could you better explain and justify the choice of this method instead or other more classical methods (PDF, MAT, PSL…). More references on the Bayesien reconstruction method are required: where this method has been tested and applied? For which time periods?

Results and Discussion

- Transfer function performance
    - The fig 3 (observed/reconstructed) is not discussed at all in the text. More sentences are needed to comment the performance of each method, for example: some high precipitation values are clearly underestimated with the WAPLS: why? May be these samples are biased by human impact and could be considered as outliers and then removed from the dataset.
    - Line 20-23: The R2 for PANN is always lower than the R2 for temperature; I don't understand why the authors compare their R2 with R2 in China; more European or Mediterranean calibrations are available (check the bibliography: Bordon et al., 2009…). It will also be important to test the spatial autocorrelation (see the papers by Telford and Birks, 2009 and others), to evaluate the performance of the models, did you do it?
    - Line 25-26: samples biased by human impact must be removed from the dataset.
- Evaluation of the reconstructions
    - Line 5: please reformulate: some differences in the levels of reconstructed Pann values

- Precipitation trends
    - Replace Late Pleistocene by Lateglacial
    - Line 19, 21: Pann is not a record, it's a reconstructed value, clarify
    - It's hard to see on the figures 4, 5 the climate patterns discussed in the text. For example: line 21 " show an increasing trend between 14500 and 14250 cal BP". The scale of the figure is not adapted to follow the discussion. Please correct.
    - P. 8, line 3: I don't agree with the author's interpretation: the pattern at Q de la Sierra is not stable and is not in agreement with a relatively stable rainfall pattern in northern Iberian Peninsula during the younger dryas.
      Line 24: The comparison with the lake levels is hard to follow; Estanya lake DOESN'T reflect large climate changes during Younger Dryas (fig 7): correct it.

- P.9, line 2 "… lower during the period 12.900 to 11.700 cal BP": to be nuanced; it depends where in Spain: in Villarquemado, Fuentillo and Padul, the lake levels were high during the Younger Dryas.
- P9, lines24-26, need to better explain the reasons of differences between proxies reconstructions: precipitation seasonality…, check the papers by Magny et al 2013 and others
- 8.2 ka event: discussion on this major event is too short, check the references on the 8.2 ka event (Magny et al 2003, 2013…) to improve the discussion
- P11: I don't understand why the authors discuss the paper by Samartin et al, (2017) which is based on two chironomids records located at high elevation sites in northern Apennines; this paper focus on the reconstruction of the temperature of July while the aim of the paper of Illvonen is to reconstruct precipitations. Moreover, the Tjuly reconstructed values are based on a modern chironomids dataset with only samples from Scandinavian and Alps; it's not comparable to Mediterranean taxa! So please avoid this discussion, or provide Tjuly reconstruction from pollen and discuss it.
- P11, line 4: I don't see where are the "reasons explained earlier"
- P11, line 14: do you take into account *Pteridium* in our dataset? I don't think so, so may be exclude these samples.

- A lot of references on past climate changes in the Mediterranean area are missing; they are needed to improve the discussion: Combourieu-Nebout et al., 2013; Bini et al., 2019; Peyron et al., 2013, Magny et al., 2013, Moreno et al., 2017…)

  Figures
  - figure1: the two regions Eurosiberian and Mediterranean must be indicated on the map
  - fig2: may be better in supplementary material
  - fig 3: not discussed in the text, to be done; check the outliers and remove it if they are linked to human impact
  - fig 4 and 5: to discuss the climate trends, you have to you trace the figures in **anomalies** (differences between past and 0k value) to avoid altitude bias.
    For clarity, I strongly recommend to the authors to merge the figures 4 and 5, and to put on the same graphs the curves obtained with both methods (all in anomalies).
    The different chronozons must appeared on the figures: GS1(or Y Dryas), Holocene to help to follow the discussion.
  - fig 6: I'm not convinced by the comparison of precipitations and temperatures; Would be better to compare with available Pann reconstructions in Mediterranean area (Tarroso et al, 2016; Peyron et al 2011, 2013, Mauri…), speleothems data…
  - fig 7: The different chronozons must appeared on the figures: GS1(or Y Dryas),in dot…

I think that this paper is interesting in terms of results and discussion; however a discussion more detailed and a regional comparison must be added to better support the interpretation before acceptation.