Clim. Past Discuss.,
https://doi.org/10.5194/cp-2019-33-RC2, 2019

Climate of the Past Discussions

Interactive comment on "Quantitative reconstruction of precipitation changes in the Iberian Peninsula during the Late Pleistocene and the Holocene" *by* Liisa Ilvonen et al.

Anonymous Referee #2

Received and published: 1 May 2019

Summary: The authors reconstruct mean annual precipitation for the last 15,000 years based on fossil pollen data from 7 sites in Spain. The reconstruction is based on 2 methods (WA-PLS and a Bayesian method) using a modern calibration dataset of 236 samples. They find drier conditions during the Late Glacial and late Holocene, and wetter conditions during the early-mid Holocene. This is consistent with lake level reconstructions.

The paper is well written and the methodology and approach is well established. The results are not so surprising since they have already been shown by more comprehensive pollen-based climate reconstructions by Mauri et al 2015 using MAT and Tarroso

et al 2016 using PDF method. The site of Quintanar de la Sierra was also the subject of a MAT pollen-climate reconstruction as long ago as 1997 by Penalba et al, but the authors do not appear to be aware of this. The novelty is in the use of the WA-PLS and Bayesian pollen-climate methods, which is a welcome addition to work in the region, although the calibration dataset and number of fossil pollen sites is small.

The main issues are:

1. Recognition and comparison with earlier work. This is not the first pollen-based precipitation reconstruction for the region. The authors briefly mention some of these studies, but do not make a comparison despite the fact the data is publicly available. For instance Mauri et al 2015 (https://www.ncdc.noaa.gov/paleo-search/study/18317) and Tarroso et al 2016 (doi:10.5194/cp-12-1137-2016-supplement) are both much more comprehensive in terms of the number of sites that they use in Iberia. It would be useful if the authors could compare their own work with these other studies and provide a figure which does so. On an individual site basis, it is strange that the authors also make no comment about the work of Penalba et al 1997 (doi:10.1006/qres.1997.1922) at the site of Quintanar de la Sierra. This is a much earlier pollen-based reconstruction which looks very much the same as their own reconstruction from this site.

2. Training set. The calibration data set is very small (236 samples). The study assumes that the entire Late Glacial to Holocene climate and related vegetation changes are to be found entirely in this sub-set of the current climate and vegetation of Spain. There are many thousands more modern pollen samples available from the European Modern Pollen Database, not only for Spain, but for adjoining areas which may offer more appropriate climate and vegetation analogues for the fossil samples. The authors actually cite the EMPD (Davis et al 2013), but do not provide an explanation as to why the rest of this data was not used.

3. Human impact. The authors appear to fundamentally contradict themselves when on the one hand they say how good the performance of the transfer function is when evalu-

ated using modern pollen samples, and on the other hand they say that reconstructions based on pollen samples from the late Holocene are biased because of human impact. Presumably the modern landscape is most likely the most human impacted of the entire Holocene, so if the transfer function gives good results for the modern, why is it unable to cope with the late Holocene? The authors make a number of broad and unsubstantiated statements about biases in pollen-climate reconstructions due to human impact. These need to be toned down and contextualised with actual evidence of bias, not just evidence of human impact. The authors also need to recognise that there are many different methods and approaches to pollen-climate reconstruction, some more sensitive than others to human impact.

4. Data transparency. As a minimum, the authors need to provide enough information so that someone else could independently reproduce their study. Unfortunately, this is not currently possible with the level of information provided. The necessary information is highlighted in the detailed comments below. This information needs to be included in the supplementary data, and it includes the taxa assignments, the sites/samples and chronologies taken from the EPD and EMPD, as well as other pollen and chronological data. It would be even better if the primary data that is not already in the public domain is also included in the supplementary, or made available via a public database.

5. Evaluation. The authors compare their pollen-based precipitation reconstructions with lake level records, but the records are few, fragmentary, and much of it involves qualitative discussion. As it stands this study is relatively weak in the sense that there have already been previous pollen-based precipitation reconstructions for the region that are in many ways more comprehensive. A good way to strengthen the paper would be to provide a more comprehensive review of precipitation records in general, including lake level data. Morellon et al (2018) provide a good example of lake level synthesis which the authors partially reproduce in figure 70 p30. Morellon et al (2018) only look at the period 8-13k, but the authors here could extend this to encompass the entire Holocene. There are also plenty of lake level reconstruc-

tions that Morellon et al (2018) do not include such as Sanchez Goni, Las Pardillas (1999) (doi: 10.1191/095968399671230625, table 4), Davis et al, Los Monegros (doi:10.1016/j.quascirev.2007.04.007, figure 8) and Reed et al, Laguna Medina (doi: 10.1191/09596830195735, figure 7).

Detailed comments:

P2, line 5-6: Mediterranean climate is specifically characterized by wet winters and dry (growing season) summers, not 'dry and wet seasons'

P2, line 21: what is a 'synthetic climate reconstruction'?

P2, line 22: there are lots of other studies you could mention..

P3, line 30: Why only 236 samples?

P3, line 31: Please provide as a minimum the condensed list of taxa that were used in the transfer function, and preferably also the full taxa list showing how all taxa were assigned to the condensed list used in the transfer function. In addition, please explain how you chose the particular taxa in the condensed taxa list used in the transfer function.

P4, line 12-13: Please provide a full list of the 236 modern pollen samples that were used, their location (lat, long and elevation) and the rainfall values assigned to these locations. A maximum of 1327mm/year seems quite low considering how wet the temperate parts of Iberia can get, and also considering the need for representative analogues for the temperate vegetation that dominated many Mediterranean areas in the early-mid Holocene.

P4, line 12-13: If samples were taken from the European Modern Pollen Database, or included in it, please include the full EMPD identity reference codes/numbers so that all of the samples can be identified. For any other samples it would be preferable if the pollen data was made available in the supplementary information or via submission to a public database or repository.

P4, line 15-19: Both Mauri et al 2015 and Tarroso et al 2016 used many more pollen sites from the region to reconstruct precipitation, since many more are available from the EPD. Can you explain why and how you chose your sites, and the basis of your reasoning for excluding the ones that you did.

P4, line 15-19: Can you identify in the table of sites, or in the supplementary information, the exact entities (you need the EPD entity reference code) and chronologies (some sites have multiple chronologies or choice of control points) you used for the EPD sites so that the primary date can be identified. For the remaining sites please specify the exact source (author or Paleodiversitas) for each of the sites in the same table. Can you also specify whether any of this data has been made public, and where this data can be downloaded. For the data that is not public, it would be preferable if it was included in the supplementary information, as well as being submitted to the EPD.

P4, line 25: 'To produce chronology' please correct the grammar

P4, line 25- P5, line 5: It would be preferable if you provided the full chronological information for each of the sites, including all control points, depths of dates, the (uncalibrated) dates themselves and their uncertainties, material dated and reference codes. Also, please say if any corrections (eg for reservoir/hardwater effects) were applied. For the EPD sites, it would be a nice gesture if you also submitted your chronologies to the EPD since they are probably better than the existing chronologies.

P5, line 7-20: You mention the complexity of the response of vegetation to climate. Is annual precipitation a reliable metric? The city of Paris gets around the same annual precipitation as Barcelona (630mm/year). Is it not growing season moisture that is more important in determining the response of vegetation to precipitation in a Mediterranean climate with hot dry summers and cool wet winters? Annual precipitation does not necessarily give any idea of the amount of growing season moisture. For instance, what would happen if the annual precipitation remained the same, but all the winter rainfall fell instead in summer? Presumably the impact on the vegetation would be

dramatic.

P5, line 19-20, P6, line 8-18: The performance metrics are only as good as the dataset that you are analyzing. If the dataset is biased (for instance, it does not sample across the entire climate space) you would not necessarily discover this from these kinds of analysis. It is not really possible to compare the performance of different datasets from different regions without standardizing the many factors that influence each dataset. In summary, r2 and rmsep are just one line of evidence for determining the performance of a transfer function, since they can easily give good results with bad data.

P6, line 20-22: Be very careful about making broad unsubstantiated statements. Evidence of human impact does not mean the same as evidence of bias. There are many different methods for reconstructing climate from pollen data, some more susceptible than others. Li et al 2014 use WA-PLS, small calibration datasets and a single site example, all of which could be expected to perform poorly in areas of heavy human impact. The main conclusion of Li et al that human impact biases the pollen-based climate reconstruction is based almost entirely on correlation, or lack of, between the pollen-based temp/precip record and other records that can anyway be expected to be different because they represent different spatial scales, temporal resolutions, or represent entirely different sensors/proxies (speleothem isotopes are a combination of precipitation and temperature signals at the destination, as well as SST and isotopic ratio of the source). Li et al also don't mention the importance of the uncertainties of the pollen-climate reconstruction in making these comparisons (one of the main effects of human impact should be to increase the uncertainties if the transfer function works correctly). Again, no one is denying that human impact can be important in pollen climate reconstructions, but you need to be careful about your evidence and phrasing here.

P6, line 26-28: Many more modern surface samples are available from the EMPD from a wide variety of depositional environments (https://epdweblog.org/european-modern-pollen-database/). I should also say, as I am sure the authors also know, that the

depositional environment also changes through time. For instance, a modern peat bog/mire may have been a lake at some time in the past, so what may make sense for the present may have to be adapted through time.

P7, line 10-11: See point 5 in the opening remarks.

P7, line 17: There are many other sites from the EPD that cover this period.

P7, line 17-28: Please read the work of Penalba et al 1997 (doi:10.1006/qres.1997.1922) at the site of Quintanar de la Sierra. The pollen-climate reconstruction in this paper looks very similar to your own.

P7, line 26-28: Be careful conflating different seasons in these comments. The Chironomid reconstruction is for summer temperatures, but your reconstruction is for annual precipitation. An increase in annual precipitation may be driven by wetter winters, unrelated to warmer summers shown by the chironomids.

P10, line 17: replace 'can be also' with 'can also be'

P10, line 23-24: This is misleading. The chironomid summer temperature record from Basa de la Mora lake by Tarrats et al 2018 indicates warmer temperatures in the early Holocene relative to the mid-late Holocene, but these temperatures were either similar or cooler than the present day. See figure 5a in Tarrats et al. The authors reconstruct a modern July air temperature of around 9.5C but they reconstruct early Holocene temperatures of around 9.1C. In fact the present July temperature for the site based on the New et al 2002 climatology adjusted for altitude is 13.2C. Tarrats et al 2018 suggest that the late Holocene samples are unreliable due to human impact, so the early Holocene summer temperatures in the chironomid reconstruction (9.1C) would in fact appear to be 4C cooler than the present day climate (13.2C).

P10, line 33: As above, Tarrats et al 2018 may SAY that their reconstruction supports Samartin et al 2017, but their reconstruction does not show this. At best, their early Holocene summer temperatures are similar to the present day if you believe the late

Holocene reconstruction (Tarrats et al, figure 5a). Or at worst, the early Holocene summer temperatures are in fact substantially cooler than modern if you follow the authors recommendations and reject the late Holocene reconstruction as unreliable.

P10, line 30-34: The Mediterranean is a big place. Samartin et al 2017 provide evidence of summer warming from 2 small adjacent lakes high in the northern Italian Apennines. This is not controversial, nor does it challenge previous work, nor should it necessarily provide support for climate models. Any analysis of multiple sites in the region will show areas of warming and areas of cooling in the early-mid Holocene. This is not surprising as it is entirely consistent with what is found in virtually every other region of the world where multi-site studies of Holocene climate have ever taken place (eg North America; Viau et al doi:10.1029/2005JD006031, Arctic; Kaufmann et al doi:10.1016/j.quascirev.2015.10.021, China; Wu et al doi: 10.1007/s00382-007-0231-3, North Atlantic; De Vernal et al doi:10.1016/j.gloplacha.2006.06.023 etc). Areas of both warming and cooling are shown in the Mediterranean in the first multi-site pollen reconstructions of summer temperatures by Huntley & Prentice (1988) (doi:10.1126/science.241.4866.687), it is shown in Wu et al (2007) (doi: 10.1007/s00382-007-0231-3) it is shown in Mauri et al (2015), based on hundreds of pollen sites from across the Mediterranean, and it is even shown in SST records from across the Mediterranean (Hessler et al 10.5194/cp-10-2237-2014). The scientific basis for a cooler early-mid Holocene is based on the regional area-average temperature calculated from the temperature trend shown at tens or hundreds of sites from across all areas of the Mediterranean (Davis et al 2003, Mauri et al 2015 etc). It is nothing new to find that individual sites and small areas show warming, it is the regional area-average that is important. This regional area-average approach is also the appropriate spatial scale to make comparisons with global climate models, in which the Mediterranean is represented by just a few grid boxes, and where countries such as Italy are barely resolved at all. For Samartin et al to truly challenge this view, or support the results of the GCMs, then they need to come up with a comparable analysis at a similar spatial scale. Even during the ubiquitous greenhouse driven global warming of the last

150 years it is still possible to find individual sites and small local areas that show cooling, but we don't then pretend that the presence of these odd sites/areas represents the regional/global picture.

P10, line 1-5: Please do not make unsubstantiated statements. Why specifically do you think that pollen-based temperature reconstructions are unreliable in the region? Please cite supporting evidence (and not other studies that make similar unsubstantiated statements such as Samartin et al 2018). Pollen-based climate reconstructions reconstruct a cooler climate in the early-mid Holocene at most (but by no means all) locations largely because of the presence of temperature vegetation (eg Prentice et al 1996 Reconstructing biomes from palaeoecological data: a general method, Collins et al 2012 doi:10.1111/j.1365-2699.2012.02738.x). The idea that these reconstructions are being driven by changes in 'Mediterranean vegetation' as the authors suggest is not therefore entirely correct. In any case the authors idea that precipitation alone drives the observed mid-early Holocene vegetation change, and not temperature, has already been tested using a process-based vegetation model in diagnostic mode by Wu et al. 2007 (doi: 10.1007/s00382-007-0231-3). Wu et al. investigated what the most probable climate would have been to generate the mid-early Holocene vegetation shown at pollen sites across the Mediterranean. They show that changes in precipitation alone cannot account for the change in vegetation, and that cooler summer temperatures are still necessary to reproduce the vegetation changes shown in the pollen record. Note that this result is also entirely independent of any potential bias that the authors suggest has been caused by later human impact on the modern vegetation. The agreement between the inverse modelling approach by Wu et al, and the calibration-based approaches by Huntley et al, Mauri et al etc, suggest that contrary to the authors suggestions, human impact is not a proven cause of bias in pollen transfer functions in the Mediterranean, at least in studies such as these that use very large training sets.

P11, line 29-33: again, please do not make broad unsubstantiated statements. Why specifically do you think that pollen-based temperature reconstructions are unreliable

in the region? Please cite supporting evidence. Evidence of human impact on vegetation is not evidence that pollen-based climate reconstructions are 'strongly influenced' by human impact. This is not to deny that the problem exists, but it is important to recognize that not all pollen-based climate reconstructions are the same, and that some have been designed specifically to limit this problem. In particular, large scale studies based on hundreds of sites and training sets of thousands of samples are arguably better able to capture the underlying regional climate signal from the noise generated by localized human impact. Other strategies include grouping taxa into common plant functional groups, and in many cases samples with high uncertainties and poor analogues that can be related to human impact are in any case filtered out. Studies based on only a few or even a single site, and those that use small training sets, are much more vulnerable. However, as the authors present themselves, transfer function performance is almost always evaluated using modern pollen samples from the modern landscape which has probably the highest human impact of the entire Holocene. The fact that we get such good performance statistics says something about the robustness of these transfer functions, although as always it is important to remember that performance statistics can be misleading. Human impact introduces noise in the relationship between vegetation and climate, usually resulting in wider uncertainties rather than simply 'wrong' values, but training set size is critical here since small datasets can be very sensitive. There are also other ways to evaluate performance. For instance, as mentioned before with the Wu et al study, cooler summer temperatures in the early-mid Holocene are also reconstructed by inverse modelling methods independent of the issue of human impact in the present-day landscape. Similarly, as the authors themselves find, the late Holocene increase in aridity reconstructed in many pollen reconstructions is also supported by lower lake levels and other independent proxies.

In any case, as here, the problem is often presented the wrong way around. The training set is tested and tuned to the modern human-impacted landscape, so it is not necessarily human impact that is the problem, but a lack of human impact in the past for which we have no modern analogue (note all calibration based methods require

modern analogues, including PDF, WA-PLS etc).

P12, line 1-7: Please refer to my previous comments. There is a world of difference between a reconstruction based on a single site or sample, and that based on many hundreds of sites or samples. The authors own transfer function has been constructed and its performance and uncertainties evaluated in the present day human impacted landscape. Human impact adds noise, but it is not necessarily overwhelming noise.

P12, line 18-19: Please be precise in your terminology. When you say 'high' do you mean higher than present during the period 8-4k? This is probably true for lake levels, but where is the evidence for higher summer temperatures? In the rest of the paper you appear to dismiss the pollen based reconstructions, so you are only left with the Pyrenees chironomid reconstruction by Tarrats et al (2018) which shows either comparable to present or most likely cooler summer temperatures. Are there other published quantitative summer temperature reconstructions from Iberia that support your conclusion?

P12, line 22-23: You conclude that late-Holocene climate reconstructions are 'thus substantially biased by human impact', but your own transfer function is based on the modern landscape which is probably the most human impacted of the entire Holocene. How can you make this statement and not negate your entire study, based on the fact that you are also applying your 'human impacted' modern calibration dataset to the earlier period? I think that you need to be more careful in making broad statements of this nature. Clearly human impact has some role, and there are certainly sites where at times the record has been influenced by human impact, sometimes significantly so. However, as I have mentioned in earlier comments, evidence of human impact on the pollen record is not evidence that climate reconstructions based on pollen data are 'thus substantially biased'. There are certainly poorly designed studies that would be susceptible to human impact, and the authors study has obvious weaknesses (small number of sites, small training set), but you should be prepared to present a much stronger argument if you want to summarily dismiss all of the previous work using

pollen-climate transfer functions in the region.

P12, line 25-29: See my opening comments. Please identify the source and identity of all of the data that you have used in the study, and preferably make available as much of the primary data that is not publicly available as possible. This can be in the form of supplementary information, or better, submit it to a recognized database (EPD, Neotoma) or data respository (Pangaea, NOAA paleoclimate). If you make data available in the supplementary information, please do so in a recognized data file format (excel, comma delimited etc), and not as a word or pdf file. I recognise that the authors have already generously submitted their data in the past, but it helps if we continue to aim for maximum data transparency in science, and especially in climate science given its high public profile.

P13, line 9-12. Please acknowledge the EPD and EMPD according to the requirements of the protocol for data use (http://europeanpollendatabase.net/datapolicy/). It is really critical that you acknowledge use of these databases so that funders can see where their money is being spent. Much of the work to support these databases is done by volunteers because of the lack of funding, and the lack of acknowledgement just makes this problem worse.

P24, Figure 1: I would recommend avoiding graded scaling, and especially multiple colours for a simple graduated scale. Contour scaling using simple 1 or 2 colour shading is a much clearer way to show this kind of information on a map. Look in any climate text book.

P27 figure 4, P28 Figure 5: I cannot understand the scaling. Along the top is every 400mm and the bottom is every 500mm. Please use the same scale, and also make the tick marks clearer. Also include a vertical line to see the anomaly from the present, not just a dot for the present precipitation. What are the 'formal sub-divisions of the Holocene'? please provide a citation. The Holocene transition seems reasonable, the end of the Laurentide ice sheet around 8k seems ok, at least in Europe directly

downstream, but what is the step-wise change in the Earths climate system at 4k? This all sounds like a hangover from the days of Blytt-Sernander and subsequent efforts to shoe-horn a Scandinavian framework onto the rest of the world.

P30, figure 7: Can you not provide a more comprehensive review of lake levels and other precipitation proxies for comparison with the pollen based reconstructions? Many are described in the text but not shown here. See #5 of my opening comments. I would also mention Harrison & Digerfeldt 1993 'European lakes as palaeohydrological and palaeoclimatic indicators' (see figure 10), which is old but still appears to be relevant today.

Interactive comment on Clim. Past Discuss., https://doi.org/10.5194/cp-2019-33, 2019.