Climate of the Past Discussions

Clim. Past Discuss., referee comment RC2
https://doi.org/10.5194/cp-2022-21-RC2, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on cp-2022-21

Anonymous Referee #2

Referee comment on "Simulations of the Holocene climate in Europe using an interactive downscaling within the iLOVECLIM model (version 1.1)" by Frank Arthur et al., Clim. Past Discuss., https://doi.org/10.5194/cp-2022-21-RC2, 2022

Summary

The authors present a Holocene climate model simulation for Europe at a high-spatial resolution using dynamical (as opposed to statistical) downscaling. This regional model simulation uses inputs from the global iLOVECLIM EMIC. The authors compare this simulation with climate reconstructions and conclude that their higher-resolution model better matches the data than the lower-resolution EMIC. The main innovation is applying this method to a transient simulation that encompasses the whole Holocene and not just time-slices. I am with the authors here, in that I think that scaling down model results to the spatial scale of the proxy data is a potentially good idea, and one well worth investigating.

General comments

The premise of the paper is interesting and is clearly appropriate for publication in Climate of the Past. The main problem is the nature of the evaluation which is qualitative and anecdotal rather than being rigorously quantitative. The authors really need to include a better designed evaluation process where proxy reconstructions are compared with the model results on a site/record basis, or for a region, in the case of gridded reconstructions. Improvements in the evaluation should also be extend to comparisons between the EMIC and downscaled modeling, for instance plotting both in the time-series plots. It would also be useful to compare the iLOVECLIM EMIC that is used with other models (eg PMIP3 GCM's) to see what particular biases this particular model has over the study region.

The authors also appear to be a bit loose with their commentary. When referring to temperatures and temperature changes it is essential to state what temperature variable it is that they are referring to, for instance whether it is mean annual, winter, summer or

some other aspect of temperature. This appears to be a source of confusion throughout, and particularly when citing other studies as supporting evidence. Similarly, in the discussion of proxy data, the authors need to distinguish between studies that provide evidence from single sites, and those that provide evidence from many hundreds of sites, since they are not equivalent. Also, it is important to understand the different studies cited; In Southern Europe, Bartlein et al is essentially a synthesis of the data of Wu et al 2007 and Davis et al 2003, Brewer et al 2007 takes data from Davis et al 2003, and Mauri et al 2014/15 is an improved version of Davis et al 2003 and Brewer et al 2007, both of which it supersedes. All of these latter studies use gridded data, not site data, although the gridding itself is based on site data.

I will also just add here the importance of isostatic uplift over Scandinavia during the Holocene, which in some areas has been substantial (100m+). Data-model comparisons over Scandinavia should be treated with caution where the model uses modern topography but is compared with proxy reconstructions that were much lower in the early Holocene.

Detailed comments

67-69 (and 160-164) This is a critical point.. the regional downscaling cannot correct major errors and biases in the global model simulation, including atmospheric dynamics which have been suggested as the source of much of the data-model discrepancy over Europe during the Holocene (Mauri et al 2014). The authors should note how the iLOVECLIM model generally compares with other global model simulations (eg PMIP3), for instance if it is generally cooler/warmer or wetter/drier than average, or comparable.

77-78 The main 'sensor' area of a proxy-based climate reconstruction is rarely greater than ~20 km radius for pollen and can be as small as a couple of hectares for lake-based proxies such as chironomids. It therefore makes sense to undertake data-model comparisons at comparable spatial scales (see my opening comments about improving the data-model comparison).

101-106 This is misleading. While Brewer et al does suggest that climate models can simulate cooler temperatures over Southern Europe and the Mediterranean during the mid-Holocene, this is only in WINTER and the signal is very weak. In contrast, reconstructions of SUMMER temperatures are much cooler than the models, which all show warmer summer temperatures (ie not even the same sign). This is not discussed by Brewer et al, but is clearly shown in the more recent reconstruction by Mauri et al 2014. Both Brewer et al (who uses the data from Davis et al 2003) and Mauri et al use pollen data, but the problem with cooler summer temperatures is also shown in SST reconstructions for the Mediterranean, as shown in Hessler et al 2014 (doi:10.5194/cp-10-2237-2014) figure 4. In fact, the data-model discrepancy shown in Mauri et al 2014 is a very good justification for the authors to have undertaken their study.

119-121 I don't really understand why the authors have chosen specific areas (and variables, eg precipitation) where they then say they don't actually have proxy records, if their stated aim is to make comparisons with proxy records. It seems that they did the model analysis first, and then looked for proxy records afterwards.

194-195 The authors use a pre-industrial climate baseline to calculate anomalies to compare with climate reconstructions. I hope that the authors are aware that anomalies shown in almost all proxy-based reconstructions are based on a modern baseline (apart from for instance Davis et al 2003 used in Brewer et al 2007, and Mauri et al 2014/5 that use a pre-industrial baseline of ~1850)

234+ Please be very careful, do not use the unspecified term 'temperature/s'. Please always state if this is annual, seasonal (JJA, DJF) etc. The authors appear to be conflating winter (Brewer et al 2007) and annual temperatures (Wu et al 2007) in the data, while the temperatures you are referring to in the model results are unspecified.

246 Again this is misleading. Brewer et al only considered winter temperatures. Better to refer to Mauri et al 2014/15 which is a more recent and more comprehensive study that includes summer winter and annual temperatures (and precipitation).

249 Wu et al uses an inverse modelling method, so represents a very different pollen-climate reconstruction to the MAT method used by Davis et al (in Brewer et al) and Mauri et al, although both show the similar results (see Davis 2017 https://doi.org/10.22498/pages.25.3.16). Note also that Wu et al is for individual sites, while Brewer et al use a gridded reconstruction where the site data has been interpolated onto a 1 degree spatial grid. There are also considerably more sites in Brewer et al than in the Wu et al reconstruction, while the sites used in Wu et al are poorly dated, use truncated taxa assemblages (a lot of data is from Huntley and Birks 1983), and have large uncertainties.

258 Fig.3 What aspect of temperature is the figure showing? Mean annual, summer, winter etc ? please specify

264+ Again, as with temperature, please specify what aspect of precipitation you are talking about.. I presume its mean annual (units are in mm/yr), but please state this clearly at the start

280-288 This is interesting, if not surprising. A question arising from this would be if the downscaling simply spatially redistributes the average precipitation of the EMIC grid box, or does it potentially increase/decrease the average precipitation that would occur in the EMIC grid box?

323 Fig. 5. & 371 Fig. 6, Section 4.1.1 Why have the authors chosen not to do a data-model comparison for the precipitation time-series for the Alps, Scandes etc? This would be very straight forward using the data from Mauri et al 2015 which is freely available. Also, why not show the EMIC result for the same spatial areas? This would illustrate the difference between the EMIC and the downscaling (and data)?

399-413 The Furlanetto et al paper consists of only one site from the Alps, the Mauri et al analysis consists of hundreds of sites from the Alps (this has been gridded, but there is also the underlying site data that could be used). Why did you pick only this particular study? In any case, it would be useful to plot the Furlanetto et al precipitation reconstruction against the model result (both high and low resolution) so that the reader can see for themselves. It is also notable that the authors identify the strong spatial variance of the precipitation signal, but compare this single proxy site with the average precipitation of the entire Alps. Surely it makes more sense to compare the proxy record with the nearest point in the model grid?

415-429 Again, as with the Alps, it would be better to directly show the proxy reconstructions plotted against the model result, and even better to show this at the model grid point closest to the site (or interpolated to the site location). If the authors really want to compare using the entire Scandes region, then at least compare against the same area using the Mauri et al 2015 gridded data, since this is designed to avoid spatial sampling bias associated with simply averaging site records together. It is also important to note that (presumably) the model uses modern topography and does not take into account the substantial changes in elevation that has occurred during the Holocene due to isostatic uplift. This is important when comparing with proxy records that actually include this isostatic change (See Mauri et al 2015).

431-435 Refer to Mauri et al 2015, there are quite a few quantitative precipitation records from this region. I am not sure why the authors say otherwise unless they want to exclude the Mauri et al data for some reason. There are also qualitative bog surface wetness records that may show trends eg Anderson 2008 (doi 10.1111/j.1502-3885.1998.tb00880.x)

448-465 The authors are conflating proxy reconstructions here across all kinds of spatial scales, some from individual sites, some based on the synthesis of large numbers of sites to represent individual regions, and some where the site records are projected onto a spatial grid. Again, it would be better if at least some of these records were compared explicitly with the model (ie one plotted over the other) rather than resorting to a rather vague ‘one thing looks like another thing’ statement, which is open to interpretation. This is particularly important because the ability to compare model and proxy record at the scale of the proxy site is supposed to be one of the main advantages of the model downscaling that the authors are proposing. Including the results of the EMIC in the same way would also help demonstrate this.

475 Southern Europe in the mid-Holocene in the PMIP2 simulations is warmer not cooler (only winter is a little cooler in the far east) and with little change in precipitation. This is

shown in detail in Mauri et al 2014.

475-478 Please do not write about climate model results as if they are some kind of reality. For instance "we *can infer from their work that southern Europe was wetter and cooler*." Should read something like "*eg we can infer from their work that southern Europe was wetter and cooler in PMIP2 model simulations*."

475-482 The PMIP2 results encompass a large number of different models, each sometimes showing quite different results. Are you talking about individual PMIP2 models, the ensemble mean or something else? Please be more specific. Also, Braconnot et al 2007 does not show any detail for Southern Europe or the Mediterranean (but is shown in detail in Mauri et al 2014) so I am not sure how the authors are making their comparison unless they have been plotting the data separately (it would be great to actually show this). In any case PMIP2 has been superseded by PMIP3.

485 Grammar needs correcting: "*in the change pattern for*"

485-495 What about temperature lapse rates? Changes in temperature lapse rates as a result of the downscaling will also lead to change in temperatures at different altitudes, perhaps better reflecting the proxy data.

496+ Conclusions- see my opening comments. The study needs a more rigorous approach to the data-model and model-model comparison.