Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2017-254-RC1, 2017

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "Comparison of spatial downscaling methods of general circulation models to study climate variability during the Last Glacial Maximum" *by* Guillaume Latombe et al.

Anonymous Referee #1

Received and published: 6 December 2017

Review for: Comparison of spatial downscaling methods of general circulation models to study climate variability during the Last Glacial Maximum

This study applies a Generalized Additive Model to statistically downscale precipitation and temperature over Europe during the Last Glacial Maximum. It specifically evaluates the effect of different interpolation schemes (bilinear, bicubic and kriging) to the application of a previously used downscaling method (Vrac et al. 2007 as cited in the manuscript). I believe this manuscript could be accepted subject to revisions concerning the following issues.

The first issue involves the coarse scale GCM predictor variables used by the down-

Printer-friendly version

Discussion paper

scaling model. Using a single GCM (IPSL-CM5A-LR) to calibrate the model and generate simulations is problematic as it leaves the analysis subject to the biases of that individual model (biases identified in “European temperatures in CMIP5: origins of present-day biases and future uncertainties” for example). This GCM has a larger than average climate sensitivity relative to the CMIP5 ensemble and it’s response to significantly reduced GHG concentrations may be similarly different from other GCMs. Calibrating a single GCM over a 30-year period should eliminate any biases due to inter-annual or decadal variability but could still influenced by lower frequency modes of variability. Magnitudes of temperature and precipitation in paleoclimate simulations could be amplified or diminished depending on whether the model was fitted in a generally cooler or warmer phase of low-frequency variability. Using an ensemble of models generally limits this effect as well. If only one GCM is feasible, then its characteristics and limitations should be explained in more detail.

The second concern is in the results for Generalized Additive Model (Section 3). There is confusion between the text and Figures 2 and 3 about what is occurring. On Page 8 starting with lines 34-35 and continued onto next page the text states: “Simulated atmospheric temperature at sea level was lower for the LGM than for the present-day period”. Is this true? In Figure 2 the legend suggests present-day SLP is lower, while the caption suggests LGM SLP is lower (the figure legend and caption contradict each other over what the solid and dashed lines represent). Further, the domain of the spline for SLP (fitted in present-day) in Figure 3 is 1000 hPa to 1030 hPa which corresponds to the lower valued histogram in Figure 2, contradicting the text.

If the spline for SLP in Figure 3 is correct then it implies LGM SLP is described by the solid line and is higher than present-day SLP. This is (hopefully) correct because if the text and Figure 2 caption are correct, the temperature panel would imply that the LGM had high temperatures than present-day, suggesting there is something seriously wrong with the IPSL-CM5A-LR GCM! If the splines of Figure 3 are correct, then a linear extrapolation of the SLP spline into the higher SLP values of the LGM suggests

Printer-friendly version

Discussion paper

precipitation will have a strong positive response to increasing SLP. This does not seem physically realistic.

The large differences in temperature in Figure S11 between downscaled and interpolated GCM values also raise doubts about the linear extrapolation of the splines to lower temperatures. The GAM is clearly adjusting the GCM temperatures upward in the majority of the region in response to what appears to be a cold bias in the GCM shown in S12 (the order of subtraction should be specified in both figure captions to confirm this). But does that mean in the LGM the GCM has a 20 degree C cold bias and the GAM is correcting this? Or is the GAM overcompensating and generating temperatures that are too warm because the slope of the temperature spline is too low?

A useful check of the downscaling model's performance would be to simulate the years in the historical model run (1901-1950 if available, 1951-1960,1990-2005) outside the calibration period and ask how the method performs against CRU observations before attempting to employ the method in time period with substantially different atmospheric forcing conditions. I suggest repeating the figures of S12 and S16 but comparing downscaled values (for winter, summer and using the different interpolation methods) against the CRU observations. If these figures replaced Figures 5 and 8 (moving those to the supplementary figures), it would provide a better picture of the method performance.

It may be beyond the scope of this study but it would be useful to see the GAM fitted separately using proxy data from the 29 sites in past and in present to see how the splines vary between such different climate regimes and whether linear extrapolation is indeed a good assumption.

The third concern is regarding the comparison of simulated temperature and precipitation against paleo-reconstructions during the LGM. The boxplots of Figures 6 and 9 do not clearly support the claim that the downscaling method is in "good agreement"

with the reconstructed values. I suggest removing (or moving to supplementary) the bilinear and bicubic panels and instead display comparisons of annual maximum, minimum and mean values separately for the kriging simulations using proper boxplots. This would provide a clearer comparison between the actual values and allow for at least a visual comparison of the distribution of these values over the 50-year LGM period to be compared. Additionally, how do you measure model performance when the two selected proxy biomes are significantly different from one another (as occurs more often for precipitation)?

It would also be particularly useful to evaluate the performance of the GAM in replicating present variability outside the calibration period given the importance of climate variability for human population distributions. Figures S19 illustrates the differences between interpolation methods in the LGM but doesn't show whether the GAM is simulating the variability accurately. It would be useful to see SPI and STI from the GAM compared to the same values for CRU similar to S12 and S16.

Further to Figure S19, maps showing the differences between the interpolation methods, as presented in the figures before, would help illustrate the effect of the different methods more clearly. Are the differences in variability from the three methods meaningful and if so are they large enough to suggest the methods could imply different patterns of human migration?

Specific/Technical Comments:

P1 Line 17: Remove the "s" from methods in "Statistical Downscaling Methods".

P1 Line 27: In the sentence beginning with "Our results" replace "confirming" with "suggesting", add "is" before "suitable" and drop "is sound" at the end. The current sentence is too strong given the evidence presented.

P1 Line 31: Replace "their" with "the".

P8 Line 3: I am skeptical that the p-value for ACO is so low (particularly for temperature)

given the sensitivity of the GAM to ACO is so small. If the variance explained by ACO is indeed statistically significant, the splines and the AIC values would suggest it is not meaningfully significant. This is noted in later paragraphs on this page.

P8 Line 8: The inverse proportionality of temperature to elevation in the GAM spline does not itself imply that the GCM overestimates temperature at high elevation (though it likely does for the reason stated in the next sentence). It merely implies, that in the GAM if elevation increases while the other parameters are constant, then the simulated temperature is expected to decrease.

P8 Line 16: Sentence beginning with “This should...”. The curvature of the lower end of the temperature spline is not negligible so this is not necessarily a safe assumption.

P10 Line 6: Add “s” to “underestimate”.

P10 Line 30-31: Given the boundary condition issue is present for all the interpolation methods, why not reduce the applicable study area to exclude the outer regions where the downscaled values will be unreliable?

P11 Line 15: “Satisfying results” is subjective, prefer a quantifiable description of how the results compared.

P11 Line 16-16: Sentence beginning with “Elsewhere” seems misplaced here.

P11 Line 20: “Critical” is too strong a descriptor here. This study shows the choice of interpolation can reduce spatial artifacts but does not explicitly demonstrate that it alone is most responsible for the GAM accuracy.

P11 Line 31-32: “non-linear” One could have linear splines and still end up with differences due to choice of interpolation method.

P12 Line 12: “Assuming ... accurate”. This is not a good assumption and I suggest simply starting the sentence at “We conclude”.

P12 Line 17: “Reliable temperatures”? There are significant biases in the mountains

Printer-friendly version

Discussion paper

as shown in Figure S12.

P12 Line 21 starting with “This correction” to the end of the paragraph: Isn’t this further evidence that the domain of the study area should be reduced to areas with paleo proxy data and without coverage by an ice sheet?

Figure 1 and 3: Add the linearly extrapolated splines in a different colour to show how the variables would respond in regimes that occur during the LGM.

Figure 2: Correct the labelling contradiction between the legend and the caption.

Figure 3: Are the units for the precipitation spline “mm” or “mm/day”?

Figures 6 and 9: Please revise the y-axis ranges of the boxplot figures to span the actual range of data displayed (e.g. there are not any temperature values above 40C yet the plot extends beyond 60C).

Figures 10 and 11: I understand the colour scales here vary from panel to panel to highlight spatial artifacts but it makes interpreting the relative effects of the methods more difficult. I think common colour scales would be more useful given the spatial artifacts should be visible from the contours anyway.

Figures S1 through S7: There are too many individual panels within these figures and they have insufficient resolution which makes them impossible to read. I suggest presenting only the four seasons for S1, and a few representative panels from the different interpolation methods for S2 -S7 which would allow them to be presented at a readable scale.

Figures 10 and S17: Please revise the red-black colour schemes to something analogous to the other figures. The large magnitude darker colours obscure the contours and make large areas of the map seem overly homogeneous.

Figures S17 and S18: Add a note in the caption why a different land-sea mask is used in these figures relative to all of the others. I suspect it is because the Mediterranean

illustrates the differences in interpolation technique quite well. However, if these are masked out and not used for projections in the LGM it also raises the question of whether these differences are meaningful in the areas actually used in the analysis.

Figure S19: Specify this is during the LGM. “(STI and SPI values in]-1,1[)” is a typo?

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-254, 2017.