ANONYMOUS REFEREE #1

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.

*We thank the reviewer for taking the time to review our manuscript and for providing the constructive comments below.*

MAIN COMMENTS

GC1: The first issue involves the coarse scale GCM predictor variables used by the downscaling 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.

*GR1: One purpose of the downscaling method, in addition to generating data at a finer grain than generated by the GCM, is actually to correct for the potential biases of the specific GCM. Such correction is possible thanks to the GAM, which acts as a transfer function, and is calibrated using the comparison of the interpolated GCM data with the target CRU data. This point has now been emphasised in the methods (p.4, l.12-20; p.5, l.13-16), where we discuss the biases of the GCM, and in the discussion (p.13-14, l.30-4) It is true that for application of the data to specific issues, an ensemble of models, which would require calibrating a different SDM for each model to correct for the specific bias of each model, would provide more accurate predictions. However, the purpose of this paper is to introduce the method and show how to apply it to a specific model, and such application is therefore out of the present scope. To clarify this point, and in response to the first comment of reviewer 2, we clarified the objectives of the paper in the introduction (p.3-4, l.22-6)*

GC2: 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 precipitation will have a strong positive response to increasing SLP. This does not seem physically realistic.

*GR2: We thank the reviewer for point this error out. We indeed made a mistake when describing the spline of the sea-level pressure, and the interpolation occurs on the right-had extremity of the spline. Please note that the SDM consists of applying a correction to the GCM precipitation using the other atmospheric variables, i.e. to correct the potential biases of the CGM (please see also our response to the previous comment). It does not represent a causal relationship between the predictors and precipitation, and the splines cannot be interpreted separately. In other words, the positive slope of the spline indicates that, according to the comparison with the CRU data, precipitations should be higher at high SLP than they are in the GCM. This point has been clarified on p.10, l.14-16. Moreover, given the low slope of the spline at this point, and the fact that most pressure values are below ~1045 hPa, we believe this interpolation will have limited impact on the output of the SDM. The text was modified accordingly on p.10, l.17-19.*

GC3: 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?

*GR3: We agree with the reviewer that the difference in temperature in the North-Eastern end of the study area should be considered carefully (in fact, this point was discussed in the discussion, originally on p.12, l.19-28, now p.15, l.11-11, and see also p.15, l.13-21 on*

*the impact of the calibration area of the results). We believe this difference is a combination of both the underestimation of temperature by the GCM, and an overcorrection of the SDM. However, as we now clarified in the methods (p.6, l.1-8), we are focusing on downscaling the region that was occupied by human populations during the LGM, i.e. mostly Western Europe, South of the ice-sheets. We nonetheless downscaled the whole region to explore in more details the behaviour of the SDM. The paragraph in the discussion was slightly modified to improve clarity, and now reads: "Because the GCM generated reliable temperatures at coarse grain for present-day conditions, which were highly correlated with the CRU present-day temperatures, the three interpolation techniques produced similar linear splines and led to relatively similar values for this variable. The IPSL-CM5A-LR GCM is known to predict lower temperatures than observed at high latitudes in winter (Dufresne et al., 2013). This bias was indeed observed when comparing the interpolated temperature with the CRU present-day data. As a result, the spline for temperature had a shallow slope at low temperature (Fig. 1). This correction was emphasised for the LGM data generated by the GCM in winter in the North of Europe (Fig. S11), which are outside of the range of present-day temperature, and therefore relied on a linear interpolation of the spline. The large difference in temperature is therefore likely to be a combination of an underestimation of temperature, and an over-correction of the very low temperature by the SDM. However, as stated previously, we are especially interested in downscaling climate data for the region occupied by human populations during the LGM. For the purpose of studying the spatial distribution of modern human population, this overcorrection will have negligible effects, since this region was covered by an ice cap during the time of interest (consequently, no palynological or vertebrate data were available for this region), and the range of values over the whole region in the present-day data encompasses the range of values for the region where humans were present during the LGM (Figs. S2-S7).*"

GC4: 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.

*GR4: The downscaling performance was validated on the 1950-1960 period. The CRU TS v. 1.2 time series (Mitchell et al. 2004) was used, since it is based on the same methodology used for generating the 1961-1990 climatology used for the calibration and had the same 10 minutes spatial resolution. Since the objective of the work was to apply the method to LGM data for paleo-anthropological research, we decided to keep figures 5 and 8 (now combined with figures 4 and 7, as figures 5 and 7) in the main text,*

*and to add the validation figures in appendix (Figures S16-S19). The validation was also performed for the kriging technique only, since it is the technique we recommend (now more explicitly in the discussion, p.15, l.8). The results show good agreement for the average temperature and precipitation values, and some small scale variations but overall good agreement in the general spatial patterns for the variability measures.*

GC5: 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.

*GR5: We thank the reviewer for this suggestion. However, fitting a GAM over the 29 sites wields several potential issues. Because GAMs are very flexible, using only 29 points may lead to an overfitting of the GAMs, especially when using 6 variables. Moreover, no precise values were available for the past, and we had to rely on reconstructions with confidence intervals, which could be quite large, especially for the BCI technique. We therefore believe that the best way to test the agreement is by comparing the simulated temperature and precipitation with the reconstructions, as we present in Figures 6 and 8.*

GC6: 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)?

*GR6: We have clarified the meaning of the reconstruction ranges from the two methods, which must be interpreted differently (p.8, l.20-24). Given that these reconstructions were generated in independent studies, the corresponding ranges are not directly equivalent to our temperature and precipitation ranges, which are the mean, minimum and maximum values over the 50 downscaled year. The BCI provides a minimum and maximum value over the whole zonobiome, and therefore generates wide ranges. Moreover, that means that simulated temperature and precipitation values close to the extremes of the BCI ranges is expected. The reconstructions by Wu et al. (2007) provide mean temperature of the coldest and warmest months and the ranges are therefore much smaller. However, these comparisons still offer valuable insights to evaluate our data. The meaning of the overlap between the simulated and reconstructed range has also been clarified in the results (p.11, l.10-16 and p.12, l.2-7). Note also that we*

*changed figures 6 and 8 following recommendations from reviewer 2 and now use scatterplots rather than boxplots.*

GC7: 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?

*GR7: As explained in our response to comment 4, the downscaling performance was validated on the 1950-1960 period on the CRU TS v. 1.2 time series. The results show some small scale variations but overall good agreement in the general spatial patterns for the variability measures. Note however that the CRU time series also relies on interpolation techniques (thin-plate smoothing splines) on irregularly space weather stations, and is therefore likely to suffer from its own specific biases. Small scale differences should therefore be interpreted with caution.*

*All figures presenting the results of the three interpolation techniques were modified to present only kriging and differences between kriging and the other two techniques, as suggested by the specific comment 19 of reviewer 2, to improve clarity and better show the differences between techniques. It is of course difficult to precisely assess the impact of the differences between the three techniques on patterns of human migration, but given that other studies (Burke et al.2014, 2017) found that variability is a key factor governing human distributions, we recommend using the technique providing the best results.*

SPECIFIC/TECHNICAL COMMENTS:

SC1: P1 Line 17: Remove the "s" from methods in "Statistical Downscaling Methods".

*SR1: This has been corrected.*

SC2: 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.

*SR2: This sentence has been rewritten.*

SC3: P1 Line 31: Replace “their” with “the”.

*SR3: This has been corrected.*

SC4: 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.

*SR4: The fact that the spline is significant is not surprising, given the number of points used for the calibration (the p-values are very sensitive to the size of the dataset, and p-values have been criticised for this, but since they are still the norm, we reported them nonetheless). The flat spline indicates that the effect size of ACO is small, which is a different matter.*

SC5: 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.

*SR5: This sentence has been rewritten as: “…which means that the coarse-grain temperatures generated by the GCM are higher than observed at fine grain at high elevations”.*

SC6: 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.

*SR6: This sentence was removed, and it now reads: “However, most temperature values in the sites where human presence has been observed during the LGM are within the range of present-day temperature, and the few remaining values are within 10 degrees of the minimum temperature. For very low temperatures during the LGM, the SDM outputs should be interpreted carefully, as we discuss below.”*

SC7: P10 Line 6: Add “s” to “underestimate”.

*SR7: This has been corrected.*

SC8: 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?

*SR8: The size of the area to exclude would vary with the interpolation technique, and is difficult to estimate. For transparency, we therefore decided to provide the full results. Moreover, considering such boundary conditions provides additional details on the differences between the interpolation techniques, which is one purpose of the present work. We added some sentences in the methods (p.6, l.3-8 ) to clarify these points.*

SC9: P11 Line 15: "Satisfying results" is subjective, prefer a quantifiable description of how the results compared.

*SR9: This sentence was reformulated as: "the method generated results falling within the computed confidence intervals"*

SC10: P11 Line 16-16: Sentence beginning with "Elsewhere" seems misplaced here.

*SR10: This sentence was reformulated and the new text reads as: "In a separate study, we were then able to test a suite of environmental predictors and demonstrate that climate variability is a key factor governing the spatial distribution of prehistoric human populations during the LGM (Burke et al. 2014, 2017)."*

SC11: 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.

*SR11: This sentence was reformulated as: "The interpolation technique used in the SDM had a major impact for the spatial patterns of climate variability."*

SC12: P11 Line 31-32: "non-linear" One could have linear splines and still end up with differences due to choice of interpolation method.

*SR12: This sentence was rewritten for clarification. It now reads: "The splines for these variables are non-linear and may exacerbate the differences between the bicubic interpolation and the other two techniques."*

SC13: P12 Line 12: "Assuming ... accurate". This is not a good assumption and I suggest simply starting the sentence at "We conclude".

*SR13: The beginning of the sentence was removed as suggested.*

SC14: P12 Line 17: "Reliable temperatures"? There are significant biases in the mountains as shown in Figure S12.

*SR14: After rewriting and clarifying the discussion, this sentence does not exist anymore.*

SC15: 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?

*SR15: The data used to calibrate the SDM must be a compromise between representativity and specificity compared to the area to downscale. In other words, using a region that would only cover the paleo proxys would likely not allow to have representative values for the different climate variables, and using a region that would be too wide would not allow to capture small scale variations. We added details about this point in the methods (p.6, l.1-8) and in the discussion (p.15, l13-21). We also specified that we are nonetheless downscaling the whole region to provide a more complete understanding of how the SDM operates, making clear that using the calibration region presented here is not recommended for downscaling North-East Europe.*

SC16: 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.

*SR16: Unfortunately, this feature is not available in the mgcv package in R, and we did not manage to add the linear extrapolations on the splines. However, we added the range of values for the 12 months over the 50 years during the LGM on the spline and histogram figures to improve clarity.*

SC17: Figure 2: Correct the labelling contradiction between the legend and the caption.

*SR17: The legend has been corrected.*

SC18: Figure 3: Are the units for the precipitation spline “mm” or “mm/day”?

*SR18: The units have been changed to mm/day.*

SC19: 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).

*SR19: Please note that we changed figures 6 and 8 following recommendations from reviewer 2 and now use scatterplots rather than boxplots.*

SC20: 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.

*SR20: The colour scale was changed from blue to dark red for temperature. Note that instead of showing all results, we now only show the maps of variability for the kriging technique, and show the difference between kriging and the other 2 interpolation techniques for concision and clarity, as recommended by reviewer 2.*

SC21: 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.

*SR21: The figures were modified to only show the 4 seasons, and the orientation of the page was changed to landscape to enable better readability of the figures.*

SC22: 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.

*SR22: The colour scale was changed from blue to dark red for temperature. As for the downscaled figures, we now only show the maps of variability for the kriging technique, and show the difference between kriging and the other 2 interpolation techniques for concision and clarity, as recommended by reviewer 2.*

SC23: 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.

*SR23: No land-sea mask had originally been used here because the interpolations are applied before applying the mask in the SDM. However, we agree that this was not coherent, and the mask has been applied to these figures for consistency.*

SC24: Figure S19: Specify this is during the LGM. “(STI and SPI values in ]-1,1[)” is a typo?

*SR24: Following comments from reviewer 2, the parts of the manuscript referring to the STI and SPI indices have been removed.*

ANONYMOUS REFEREE #2

This manuscript deals with the application of a downscaling technique combining interpolation (through three techniques) and General Additive Models (GAMs), over Western Europe during the Last Glacial maximum (LGM). Results are compared to site specific climate proxys from pollen and vertebrate remains data. Its seems well within the scope of Geoscientific Model Development, and deals with the relevant topic of developing statistical downscaling tools that may be used in very different climates like the LGM. The manuscript needs in my opinion some tightening of the objectives, some work on the clarity of the text and take-home messages, as well as some additional simulation analysis. I detail below these few main comments, together with many specific ones. I can therefore recommend publication of the manuscript only once all these comments are addressed.

*We thank the reviewer for these nice comments and for the constructive review he provided.*

MAIN COMMENTS

GC1. It is not clear from the start (and down to the choice of figures) what are the objectives of this manuscript. Is it the comparison of downscaling methods (i.e. through different interpolation techniques)? Is it the adequate simulation of reconstructed climate proxy data? Is the target location the whole Europe or only the proxy specific sites? All these questions should be answered from the beginning of the manuscript. As they are currently not answered, the organization of the manuscript and the choice of figures are indecisive (see specific comments below).

*GR1. We rewrote the last paragraph of the introduction (p.3-4, l.22-6) to clarify the objectives of the manuscript. As we now state explicitly, we are assessing and refining (comparing the 3 interpolation techniques) the capacity of an SDM method based on a Generalised Additive Model originally designed for the downscaling of climatology data to downscale time series, with a special interest in sites where prehistoric human presence has been recorded. We therefore seek to obtain a good accuracy for the results, while exploring the limitations of the application and acknowledging that results may be improved, for example by using an ensemble of models, as suggested by reviewer 1. We also modified and added some contents in the first paragraphs of the discussion (p.13, l.22-29) to clarify these last points.*

C2. The simulation set-up clearly lacks some present-day validation, as already pointed out by reviewer #1. This would hopefully help disentangling errors/biases from the interpolation, GAM models, and the driving GCM (see specific comments below).

*GR2. The downscaling performance was validated on the 1950-1960 period. The CRU TS v. 1.2 time series was used, since it is based on the same methodology used for generating the 1961-1990 climatology used for the calibration and had the same 10 minutes spatial resolution. Since the objective of the work was to apply the method to LGM data for paleo-anthropological research, we decided to keep figures 5 and 8 (now combined with figures 4 and 7, as figures 5 and 7) in the main text, and to add the validation figures in appendix (Figures S16-S19). The validation was also performed for the kriging technique only, since it is the technique we recommend. The results show good agreement for the average temperature and precipitation values, and some small scale variations but overall good agreement in the general spatial patterns for the variability measures.*

GC3. Another consequence of the first main comment above is that a large number of supplementary figures are commented in the main text, which is quite frustrating for the reader. The organization of figures (and associated text) should definitely be redesigned (see specific comments below).

*GR3. The figures have been re-designed and re-arranged based on the specific comments below. Given the nature of the work, we had to provide quite a number of figures in supplementary material to allow the reader to investigate some subtleties of the work in more details, while keeping the number of figures acceptable in the main text. Some figures were combined, which should increase the clarity of the manuscript. We also clarified the objectives of the present study, and we think that the current arrangement of the figures is consistent with the logic.*

GC4. In relation to the second main comment above, there is little uncertainty discussed in the manuscript, be it a result of the short calibration period for GAMs or from another source like using a single GCM. This should definitely be taken on by the authors for the manuscript.

*GR4. The biases of the GCM, which can influence the SDM have now been specified in the Methods (p. 4, l.12-20). In addition, we added a paragraph in the Discussion (p.13-15, l.13-21) in which we discuss in more details some uncertainties related to our results and make recommendation for dealing with them.*

SPECIFIC COMMENTS

SR1. P3L1: Please define “taphonomic”

*SR1. Taphonomic has been defined in the parenthesis following the word. It now reads “(i.e biases in the fossil record, such as pollen preservation, location of archaeological sites, etc.)”.*

SC2. P3L23-25: The length of the two GCM simulations is not clear here

*SR2. It is now specified that the climate data corresponds to the average of the 1961-1990 period, by contrast with the 50 years time series.*

SC3. P4L29: The reference used here for the mgcv R package is not one of those recommended in the citation info of the package. Please correct this.

*SR3. The reference has been changed to Wood (2011), as indicated in the citation information of the mgcv package.*

SC4. P5L1-2: Is there actually a theoretical reason for the requirement of the same scale? I fully understand the advantage of using e.g. downscaled precipitation as a predictor for local precipitation, but when considering other predictors like SLP, the most informative scale for local precipitation my clearly not be the local scale, but a larger domain shifted in the direction of the prevailing winds (at least in western Europe). This would open quite different approaches for performing this kind of studies that would not require the interpolation step. But this may lead to difficulties given the change in land/sea mask and the presence of ice caps when considering LGM simulations. I would appreciate a comment on that.

*SR4. We agree with the reviewer that the extents of the region used for calibration and for the downscaling are important to consider. We now explain in more details in the methods (p.6, l.1-8) that we are especially interested in downscaling the region of Europe where human presence has been observed during the LGM (i.e. South of the ice-sheets), which is why the calibration region encompasses North-East Europe, which has low temperature during the present-day period. However, using a region that is too big would not allow to capture the small-scale variations. We specify that we applied the downscaling on the whole area to better explore how the SDM performs, and added details in the discussion (p.14, l.3-4 and see also p.15, l.16-20) on the risks of using it to downscale North-Europe during the LGM.*

SC5. P5L13-15: This starts to be confusing in terms of data. I believe that (for any location) not only the monthly regime (i.e. 12 values only) is used, but the whole 31-year monthly time series. Please be more specific.

*SR5. The climatology, i.e. the 30-years average resulting in 12 values for each cell, was indeed used to calibrate the SDM, and then applied to downscale the time series. The climatology was used because the GCM is not precise enough to simulate the temperature and precipitation of a given month in a specific year, which would be require to calibrate the GAM on a time series. However, a GCM can generate temporal patterns for these variables, and we therefore tested the potential of applying an SDM calibrated on a climatology to a time series generated by the same model. These points*

*have been specified in the introduction as we clarified the objectives of the work (p.3, l.21)(cf General Comment 1), and in the methods (p.5, l.29-31).*

SC6. FigureS1: This figure is not readable at all. Same for Figures S2 to S7. I would strongly recommend finding a way to make them actually useful.

*SR6. The figures were modified to only show the 4 seasons, and the orientation of the page was changed to landscape to enable better readability of the figures.*

SC7. P528-P6L3: "Aco" should be defined mathematically in the text without having to look into Vrac et al. (2007). There is no need to define "Dco" if not used, apart maybe from writing that it is highly correlated to "Aco".

*SR7. The mathematical formula of Aco is now provided in Equation 2.*

SC8. P6L16: There should be a reference here to Table S1.

*SR8. The reference to Table S1 has been added.*

SC9. Section 2.5: There should be two additional subsections on the interpolation/downscaling for the present-day reference period (1961-1990) and for a present-day validation period (see Main comment above and comments from reviewer 1).

*SR9. A section (2.7) describing the present-day validation outside of the calibration period was added. Please note that we did not include the downscaling of the 1961-1990 climatology, because it would be redundant with the validation on present day data, and would add an unnecessary additional number of figures. Figures S10 and S13 nonetheless compare the interpolated and CRU data for 1961-1990 to discuss the correction performed by the GAM (p.11, l.2-4; p.11, l.27-29), but we considered that this did not deserve a full subsection, that would complexify an already long article.*

SC10. P6L24: "Extremes" is a much too strong word here. This set-up (length of time series and temporal resolution) prevents assessing extremes.

*SR10. "extremes" has been removed from the text.*

SC11. Figure S8: This map should definitely be included in the main text, because this critically shows where to look in European map results. It might also be relevant to systematically indicate these locations in the results maps (depending on their size).

*SR11. Figure S8 has been added to the main article and is now Figure 1. The locations were not indicated on the other figures, because they already contain a lot of*

*information, such as the contour line, and adding the site locations would impair their readability.*

SC12. P7L18-25: Results on the SPI and STI are not used at all in the manuscript (only in the Supplementary material), so please remove their description (and possible comments) from the main text.

*SR12. All parts of the manuscript referring to the SPI and STI have now been removed from the manuscript.*

SC13. P7L18-25: The description of SPI computation lacks many important details: (1) what is the climatic norm, i.e. the reference period over which the standardization is based (present-day, LGM) and why? (2) what is the chosen distribution function for monthly precipitation? (3) Is it the same everywhere in Europe? Results are quite sensitive to these issues, as clearly shown in the literature (see e.g. Wu et al., 2005; Stagge et al., 2015)

*SR13. All parts of the manuscript referring to the SPI and STI have now been removed from the manuscript.*

SC14. P7L18-25: The choice of a variability index as the number of months with SPI between -1 and 1 is actually very strange (and indeed quite irrelevant). The SPI is by definition normally distributed, so the probability of having a SPI between -1 and 1 is 68.27%, which amounts to around 410 months in 50 years (95% confidence interval: 387-432), if the reference period for fitting the precipitation distribution is the same as the computation period (which I believe is the case here, see P7L20). So the spatial pattern observed in Figure S19 is a complete artifact due to (1) the limited length of the period used for fitting the distribution, and (2) the relevance of the specific theoretical distribution used for fitting. Based on the 3 above comments, I strongly suggest removing all the analysis done with SPI/STI.

*SR14. All parts of the manuscript referring to the SPI and STI have now been removed from the manuscript.*

SC15. Figure 1: It would be great to see the range of present-day and LGM predictors in these figures in order to directly check statements made in the text P8L12-18.

*SR15. The total range of the values is represented by the range of the x-axis of the splines. This is now specified in the figure's caption. In addition, we added grey lines to show the values of temperature at the locations of the archaeological sites in the simulations for all months and 50 years.*

SC16. Figure 2: Like reviewer #1, I believe that dotted lines are for the present-day period. Please remove the wrong legend definition from the caption.

*SR16. There was indeed an error in the figure caption. The text and figure caption have been modified accordingly, indicating that dotted lines are for present and plain lines for the LGM.*

SC17. Section 3.2: As mentioned above for section 2.5, there should be an additional result section for validation the interpolation/downscaling process in a presentday period distinct from the calibration period.

*SR17. Section 3.3 has been added to present the results of the validation procedure on a present day time series outside of the period used for calibration.*

SC18. Most of results are presented at the annual time scale or for two 3-month seasons. What is then the advantage of fitting GAMs for individual months? I would expect a larger explained variance for annual or seasonal averages. I would appreciate some comments on this issue in the manuscript.

*SR18. The maps are presented combining months into seasons to condense the results and increase readability. This has now been specified in the figure captions. Applying the GAM to individual months is nontheless necessary for computing the variability of temperature and precipitation.*

SC19. Figure 4: Possible differences between the three interpolation techniques cannot be appreciated from these maps with a common colour scale, because of (1) the large spatial range, and (2) the large seasonal range. Figure 5 looks into all possible differences between the 3 techniques, making both figures relatively redundant. I would therefore recommend choosing one interpolation technique as reference (ideally the one that should be recommended in the conclusion of the manuscript) and plot (1) maps as in Figure 4 for this technique, and (2) differences from this reference with a specific colour scale, as in Figure 5. This would hopefully reduce the number of figures and make the message clearer ("we choose this technique and results with the others are not that different.")

*SR19. We thank the reviewer for this advice. Following his recommendation, we combined figures to represent the values for the kriging, and the differences between the kriging and the other 2 interpolation techniques, decreasing the total number of figures in the manuscript and the supplementary material, and making the message clearer.*

SC20. Figures S9 to S12. This is a much too high number of figures which shows that work on synthesizing results is clearly lacking. The reader should be presented two things: first, how temperature (and precipitation in a second step) is transformed by the whole downscaling process, through maps of raw, interpolated, interpolated +downscaled, and observations (CRU) in the present-day period. A similar presentation should be made for the validation period, and for the LGM period (for which CRU

observations may be replaced by the pollen and vertebrate proxys). This could be made only for the reference interpolation technique. Second, additional maps should show the differences with the two other techniques, possibly through the whole downscaling process. This would require reorganizing figures and text (P9L6-28), but for a much better clarity of the manuscript!

*SR20. The number of figures in both the main article and in the supplementary material has been reduced. Since, as we now clarify in the introduction, the purpose of the present work is to explore and refine an existing SDM method design for the downscaling of climatology data to downscale time series of simulated past climate. Since the core SDM method has been described previously (Vrac et al. 2007), we focus on presenting the results and the effect of using the different interpolations in the main article, since these represent its main contributions. Showing how things change from coarse to interpolated to downscaling is not the main focus here, because the effect of the downscaling will vary depending on the interpolation technique to compensate for the bias the interpolation may induce. Rather, comparisons with the interpolation data is used to shed light on the final results, and such maps are therefore in the appendix.*

SC21. P9L25-28, and Figure6: I am not convinced by results as presented here, as these plots are not very appropriate for identifying agreement for each site independently. I would therefore recommend trying scatterplots (with uncertainty bars as here or better uncertainty squares), with reconstructions (BCI, Wu et al. data) on the x-axis and simulations from this paper on the y-axis. The overlap of uncertainty ranges with the diagonal might better inform on the agreement of simulations with reconstructions.

*SR21. Following the reviewer's advice, we changed figures 6 and 8 to show scatterplots rather than boxplots. In addition, we clarified the differences between the ranges obtained with the reconstruction methods and the downscaling, and how to interpret the figures (p.11, l.13-16 and p.12, l.2-7).*

SC22. P9L30-P10L16: cf. comments on temperature for an additional validation period, revised figure organization, etc.

*SR22. Please see our response to previous comments GC2, SC9, SC17 and SC20. We believe that the addition of the present-day validation, the clarification of the objectives and the simplification of the figures make the article clearer and justify its current organisation.*

SC23. P10L9-11, "This is due. . . such as precipitation (Wood et al., 2004)": I don't understand why this should lead to the European-scale discrepancies noted in the previous sentence. Please make it clearer.

*SR23. This sentence has been re-written as: "This explains the discrepancies between the present-day simulations and the CRU data and, by extension, explains the adjustments performed by the SDM."*

SC24. P10L23-31: I find this paragraph a bit long, compared to other issues elsewhere that would also deserve some explanations.

*SR24. This paragraph is a bit long because it requires mathematical explanations, which can hardly be condensed without losing clarity. We clarified the other issues identified by the two reviewers in the rest of the article.*

SC25. P11L5-9: As mentioned above, please remove the SPI/STI analysis and results. C6

*SR25. All parts of the manuscript referring to the SPI and STI have now been removed from the manuscript.*

SC26. P11L25-26: So should we use kriging? Please be more specific on your conclusions about the preferred interpolation method.

*SR26. We removed "seems to" on l…, and added the following sentence at the end of the paragraph: "We therefore recommend using kriging for SDM applications based on the method presented here."*

SC27. P12L14-15, "more reliable variability": I don't understand. Please make it clearer.

*SR27. This part was rewritten as: "generates variability indices with more realistic patterns".*

SC28. P12L19-29: Well, this clearly poses the question on whether one should put confidence in GCM outputs at high latitudes (at least. . .). And for this study, this raises the following issue: should the interpolation/downscaling take place over the whole of Europe for reconstructing only a few sites located in the south of the continent. This issue should be seriously taken into account by the authors for the manuscript. Indeed, there may some biases in LGM results in the south due to present-day biases in the north via the continent-wide GAM modeling. . . I am definitely expecting comments on this potential issue.

*SR28. The reviewer is right to point out the issue of the confidence of the results for North-East Europe. We now provide additional details in the methods (p.6, l.1-8) and discussion (p.14, l.3-4 and see also p.15, l.16-20) about this point. As we now clarify, we are especially interested in modelling climate for parts of Europe occupied by human populations during the LGM, therefore excluding North-East Europe, which was covered by an ice cap. However, we applied the SDM to this region in the manuscript to explore in details potential issues with applying this method to a region with data outside of the*

*range of values used for calibration. The manuscript is therefore now more complete, not only showing how to apply the method, but also pointing out potential pitfalls.*

SC29. P13L3, “larger-scale patterning”: Could you explain and make it clearer?

*SR29. This sentence has been removed.*

SC30. P13L4-6: I am not sure this sentence is relevant here.

*SR30. The sentence has been removed.*

SC31. Table 1, “AIC weights”: this should be defined and commented in the text.

*SR31. The AIC weights have now been defined and commented in the figure caption, to avoid overloading the main text.*

Technical corrections
1. P5L25: Remove “interpolated variable”

*“interpolated variable” has been removed.*

2. P6L29-30: Redundancy of “downscaled, simulated”

*“simulated” has been removed.*

3. P7L3: Please specify that “bio-climatic indices” is abbreviated as BCI(s) in the following.

*This has been specified.*

4. P8L10: font size of “predictor”

*The font size has been changed.*

5. P11L31: “than for the temperature”

*The correction has been made.*

List of changes in the manuscript.

1. Title changed from “Comparison of spatial downscaling methods of general circulation models results to study climate variability during the Last Glacial Maximum » to « Comparison of spatial downscaling methods of general circulation model results to study climate variability during the Last Glacial Maximum”

2. Abstract changed from “The extent to which climate conditions influenced the spatial distribution of hominin populations in the past is highly debated. General Circulation Models (GCMs) and archaeological data have been used to address this issue. Most GCMs are not currently capable of simulating past surface climate conditions with sufficiently detailed spatial resolution to distinguish areas of potential hominin habitat, however. In this paper we propose a Statistical Downscaling Methods (SDM) for increasing the resolution of climate model outputs in a computationally efficient way. Our method uses a generalized additive model (GAM), calibrated over present-day data, to statistically downscale temperature and precipitation from the outputs of a GCM simulating the climate of the Last Glacial Maximum (19-23,000 BP) over Western Europe. Once the SDM is calibrated, we first interpolate the coarse-scale GCM outputs to the final resolution and then use the GAM to compute surface air temperature and precipitation levels using these interpolated GCM outputs and fine resolution geographical variables such as topography and distance from an ocean. The GAM acts as a transfer function, capturing non-linear relationships between variables at different spatial scales. We tested three different techniques for the first interpolation of GCM output: bilinear, bicubic, and kriging. The results were evaluated by comparing downscaled temperature and precipitation at local sites with paleoclimate reconstructions based on paleoclimate archives (archaeozoological and palynological data). Our results show that the simulated, downscaled temperature and precipitation values are in good agreement with paleoclimate reconstructions at local sites confirming that our method for producing fine-grained paleoclimate simulations suitable for conducting paleo-anthropological research is sound. In addition, the bilinear and bicubic interpolation techniques were shown to distort either the temporal variability or the values of the response variables, while the kriging method offers the best compromise. Since climate variability is an aspect of their environment to which human populations may have responded in the past this is an important distinction.”

to “The extent to which climate conditions influenced the spatial distribution of hominin populations in the past is highly debated. General Circulation Models (GCMs) and archaeological data have been used to address this issue. Most GCMs are not currently capable of simulating past surface climate conditions with sufficiently detailed spatial resolution to distinguish areas of potential hominin habitat, however. In this paper we propose a Statistical Downscaling Method (SDM) for increasing the resolution of climate model outputs in a computationally efficient way. Our method uses a generalized additive model (GAM), calibrated over present-day climatology data, to statistically downscale temperature and precipitation time series from the outputs of a GCM

simulating the climate of the Last Glacial Maximum (19-23,000 BP) over Western Europe. Once the SDM is calibrated, we first interpolate the coarse-scale GCM outputs to the final resolution and then use the GAM to compute surface air temperature and precipitation levels using these interpolated GCM outputs and fine resolution geographical variables such as topography and distance from an ocean. The GAM acts as a transfer function, capturing non-linear relationships between variables at different spatial scales and correcting for the GCM biases. We tested three different techniques for the first interpolation of GCM output: bilinear, bicubic, and kriging. The resulting SDMs were evaluated by comparing downscaled temperature and precipitation at local sites with paleoclimate reconstructions based on paleoclimate archives (archaeozoological and palynological data) and the impact of the interpolation technique on patterns of variability was explored. The SDM based on kriging interpolation, providing the best accuracy, was then validated on present-day data outside of the calibration period. Our results show that the downscaled temperature and precipitation values are in good agreement with paleoclimate reconstructions at local sites, and that our method for producing fine-grained paleoclimate simulations is therefore suitable for conducting paleo-anthropological research. It is nonetheless important to calibrate the GAM on a range of data encompassing the data to be downscaled. Otherwise, the SDM is likely to over-correct the coarse-grain data. In addition, the bilinear and bicubic interpolation techniques were shown to distort either the temporal variability or the values of the response variables, while the kriging method offered the best compromise. Since climate variability is an aspect of the environment to which human populations may have responded in the past the choice of interpolation technique is therefore an important consideration."

3. p.3-4, l.23-7: objectives of the paper clarified: "In this study, we explore and refine the capacity of an SDM from the transfer functions family, based on Generalized Additive Modelling (GAM), to compute temperature and precipitation time series at a fine spatial and temporal resolution for the LGM over Western Europe, south of the Fennoscandian ice-sheets. GAM is a non-parametric statistical technique that has proven reliable for capturing non-linear relationships between local- and large-scale variables and correcting the biases specific to a given GCM (e.g., Vrac et al., 2007; Levavasseur et al., 2010). The SDM used here accurately downscales the climatology (i.e., the climate averages over several decades) of temperature and precipitation generated by a GCM for the LGM when calibrated using present-day data (Vrac et al. 2007). Its ability to generate projections of the small-scale temporal patterns necessary to explain the spatial dynamics of prehistoric human populations is untested, however. In the present study, therefore, we use present-day climate data (corresponding to the average of the 1961-1990 period) extracted from the IPSL-CM5A-LR GCM (Dufresne et al., 2013) to calibrate the SDM, applying it to a 50 year-long time series of climate simulations for the LGM (Kageyama et al., 2013a, b). Interpolated values of coarse-grain variables extracted from the GCM, as well as fine-scale geographical data such as elevation and advective continentality, are used as predictors in the GAM. The result is the production of downscaled monthly values over 50 years for temperature and

precipitation, including local temporal variability in temperature and precipitation rates. In addition, we compare the impact of three different interpolation techniques (bilinear interpolation, bicubic interpolation and kriging) on the downscaling results, evaluating the resulting SDM outputs with the aid of climate proxies (palynological and archaeozoological data) and observing the impact of each technique on patterns of spatial and temporal variability in temperature and precipitation. The SDM using kriging interpolation is demonstrated to be a good compromise between computational complexity and accuracy, validated on an 11-year present-day time series distinct from the calibration period."

4. p.4, l.13-21: biases of the GCM specified: "The model performance and main biases are described in Dufresne et al. (2013) and Hourdin et al. (2013). This model version is known to have a cold (-1.4°C) bias in terms of globally averaged temperature and the bias in mean annual temperature over Europe is similar to the global value. In this low-resolution version of the model, the mid-latitude westerly winds are generally more equatorial than observed (Hourdin et al., 2013) and this is the case for the Northeast Atlantic and Europe too. The general pattern of extra-tropical precipitation over the North Atlantic and European sectors is satisfactory with respect to the annual mean (Dufresne et al., 2013). The equilibrium temperature response to a doubling CO2 is 3.59°C (Dufresne et al., 2013), which is rather high compared to other CMIP5 models. Nevertheless, the model's response to LGM forcings somewhat underestimates the cooling over Europe, as reconstructed from pollen, and is satisfactory in terms of precipitation (Kageyama et al., 2013)."

5. p.5, l.14-16: Addition: "The GAM uses these predictor variables, in addition to the original variables to be downscaled, to correct the biases of the GCM described above. This implies that a given GAM is only valid for the GCM for which it was calibrated, as it corrects its' specific biases." added.

6. p.6, l.1-7: Addition: "We are specifically interested in downscaling temperature and precipitation over Western Europe, south of the ice-sheets during the LGM, a time when archaeological data indicates that human populations contracted in size and range. The GAMs were calibrated over a wider area than the region of interest in order to avoid edge effects and include the full range of climate conditions that prevailed during the LGM, which was much colder than the present day. As a result, the calibration domain extends further North and East (where more continental conditions prevail) than the region of interest. Preliminary simulations nonetheless showed that selecting too large a calibration region averaged-out the small-scale variation we are interested in."

7. p.6, l. 29: equation of the advective continentality added.

8. p.7, l.25-28: "We also calculated indices of variability, including measures of variance and inter-annual variability for the variables of interest. The standard deviation of monthly mean temperatures for each month was calculated for the 50-year run. The

coefficient of variation of monthly mean precipitation values was calculated for the same period." Moved up due to the removal of the SPI and STI indices in section 2.7 originally.

9. section 3.1: specifications on the interpretation of the GAM splines, on p.8 l.16, 26, p.9 l.6, 14, and 17.

10. p.11, l.12: specifications added on how to interpret the pollen and BCI temperature reconstructions: "Simulated temperature ranges fall within the reconstructed ranges corresponding to the BCIs and are within the reconstructed ranges from Wu et al. (2007) for all test sites, as shown by the overlap of the red and blue error bars with the diagonal (Fig. 6a,c,e). As noted above, since the BCI reconstruction considers the minimum, mean and maximum values for each zonobiome over the whole of Western Europe (see Burke et al. 2014), some downscaled temperature values may differ from the mean, but as long as the error bars overlap with the diagonal are still in accordance with the BCI climate reconstruction."

11. section 3.2.2 Precipitation: the whole section was rewritten to improve clarity.

12. addition of section 3.3 Validation on present day data.

13. The discussion was largely rewritten and rearranged according to the reviewers' comments to improve clarity sand to incorporate the new validation results.

14. Figure 1 indicated the sites used for the temperature and precipitation reconstructions was added.

15. Figures 2-5: information on the range of data at the reconstruction sites were added on the figures, and the mistake inverting present-day and LGM curves was corrected.

16. Figures 5 and 7 combines the original figures 4 and 5 and figures 7 and 8, respectively.

17. Figures 6 and 8 use square error bars instead of the original boxplots for clarity.

18. Figures 9 and 10 present the same information as original figures 10 and 11 but using the same representation as figures 5 and 7, for clarity.

# Comparison of spatial downscaling methods of general circulation model results to study climate variability during the Last Glacial Maximum

Guillaume Latombe[1,2], Ariane Burke[3], Mathieu Vrac[4], Guillaume Levavasseur[5], Christophe Dumas[4], Masa Kageyama[4], Gilles Ramstein[4]

[1]Centre for invasion Biology, Department of Mathematical Sciences, Stellenbosch University, Matieland 7602, South Africa
[2]School of Biological Sciences, Monash University, Melbourne 3800, Australia
[3]Département d'Anthropologie, Université de Montréal, Montréal, QC, Canada
[4]Laboratoire des Sciences du Climat et de l'Environnement/Institut Pierre-Simon Laplace, Université Paris-Saclay, CE Saclay, l'Orme des Merisiers, Bât.701, Gif-sur-Yvette, France
[5]Institut Pierre Simon Laplace (IPSL), Pôle de modélisation du climat, UPMC, Paris, France

*Correspondence to*: Guillaume Latombe (latombe.guillaume@gmail.com)

**Abstract.** The extent to which climate conditions influenced the spatial distribution of hominin populations in the past is highly debated. General Circulation Models (GCMs) and archaeological data have been used to address this issue. Most GCMs are not currently capable of simulating past surface climate conditions with sufficiently detailed spatial resolution to distinguish areas of potential hominin habitat, however. In this paper we propose a Statistical Downscaling Method (SDM) for increasing the resolution of climate model outputs in a computationally efficient way. Our method uses a generalized additive model (GAM), calibrated over present-day climatology data, to statistically downscale temperature and precipitation time series from the outputs of a GCM simulating the climate of the Last Glacial Maximum (19-23,000 BP) over Western Europe. Once the SDM is calibrated, we first interpolate the coarse-scale GCM outputs to the final resolution and then use the GAM to compute surface air temperature and precipitation levels using these interpolated GCM outputs and fine resolution geographical variables such as topography and distance from an ocean. The GAM acts as a transfer function, capturing non-linear relationships between variables at different spatial scales and correcting for the GCM biases. We tested three different techniques for the first interpolation of GCM output: bilinear, bicubic, and kriging. The resulting SDMs were evaluated by comparing downscaled temperature and precipitation at local sites with paleoclimate reconstructions based on paleoclimate archives (archaeozoological and palynological data) and the impact of the interpolation technique on patterns of variability was explored. The SDM based on kriging interpolation, providing the best accuracy, was then validated on present-day data outside of the calibration period. Our results show that the downscaled temperature and precipitation values are in good agreement with paleoclimate reconstructions at local sites, and that our method for producing fine-grained paleoclimate simulations is therefore suitable for conducting paleo-anthropological research. It is nonetheless important to calibrate the GAM on a range of data encompassing the data to be downscaled. Otherwise, the SDM is likely to over-correct the coarse-grain data. In addition, the bilinear and bicubic interpolation techniques were shown to distort either the temporal variability



**Définition du style:** EndNote Bibliography Title: Police :(par défaut) Times New Roman, 10 pt
**Définition du style:** EndNote Bibliography: Police :(par défaut) Times New Roman, 10 pt
**Supprimé:** models
**Mis en forme:** Français
**Supprimé:** [1]Department
**Mis en forme:** Français
**Supprimé:** d'anthropologie
**Mis en forme:** Français
**Supprimé:** Methods
**Supprimé:** .
**Supprimé:** results
**Supprimé:** ). Our results show that the simulated,
**Supprimé:** confirming
**Supprimé:** is sound

or the values of the response variables, while the kriging method offered the best compromise. Since climate variability is an aspect of the environment to which human populations may have responded in the past the choice of interpolation technique is therefore an important consideration.

## 1 Introduction

The extent to which past climate change influenced human population dynamics during the course of prehistory is a subject of lively debate. The Last Glacial period, including Marine Isotope Stages 3 (MIS3) and 2 (MIS2) and the Last Glacial Maximum (LGM), is particularly interesting in the context of this debate (van Andel, 2003). During MIS3 the archaeological record suggests that modern human populations originating in Africa expanded into Eurasia, while Neanderthal populations gradually contracted their range before becoming extinct ~27,000 years Before Present (BP) (Serangeli and Bolus, 2008). Progressively colder and drier conditions, culminating in the LGM (19,000 - 23,000 years Before Present), are thought to have triggered further range contractions and the demographic decline of modern human populations in Europe. Climate affects the spatial behaviour of human populations directly (when conditions exceed human physiological limits) and indirectly (when it affects the distribution of resources upon which humans depend). The global climate during the Last Glacial period was characterised by a series of rapid oscillations (known as Dansgaard-Oeschger, or D-O events). These events may have acted as forcing mechanisms, affecting the demographic processes described above (e.g. Müller et al., 2011; Schmidt et al., 2012; Sepulchre et al., 2007; Jimenez-Espejo et al., 2007; Banks et al., 2013; d'Errico and Sánchez Goñi, 2003; Gamble et al., 2004; Shea, 2008). While the timing of climate events relative to large-scale patterns in the archaeological record is suggestive, the mechanisms by which climate forcing acted on human populations are still imperfectly understood. More empirical evidence is needed to validate the hypothesis that climate forcing affected human population dynamics and explore the nature and scale of the effect.

The broad demographic patterns mentioned above are the result of smaller, local-scale patterns produced by mobile groups of hunter-gatherers distributing themselves on the landscape in order to exploit available resources. The availability of these resources fluctuated both predictably (on a seasonal basis) and unpredictably (as a result of climate variability). It is by gaining an understanding of these smaller-scale patterns, ultimately, that we will be able to understand how climate forcing affects the spatial and cultural dynamics of prehistoric human populations. Previous analyses of climate forcing have used a variety of data to reconstruct the paleoclimate, such as ice-core or marine records (e.g., Bradtmöller et al., 2012; Jimenez-Espejo et al., 2007; Schmidt et al., 2012), present-day climate data (e.g., Jennings et al., 2011), and climate model simulations (e.g., Banks et al., 2008; Davies and Gollop, 2003; Sepulchre et al., 2007; Benito et al. 2017; Hughes et al. 2007; Tallavaara et al. 2015). These analyses were conducted at varying spatial resolutions, typically on the order of ~50 km x 50 km (= 2500 km$^2$). Higher-resolution climate simulations are nonetheless necessary for the quantification of climate variability at an inter-annual scale and a spatial scale which approximates the size of the catchments within which hunter-gatherer groups typically forage (~10 km from camp, or 314 km$^2$; Vita-Finzi and Higgs, 1970) making this an ideal spatial scale at which to consider the impact of climate variability on human systems.



**Supprimé:** offers
**Supprimé:** their
**Supprimé:** this
**Supprimé:** distinction
**Mis en forme:** Anglais (Canada)
**Supprimé:** progressively
**Supprimé:** The goal of this study is to develop high
**Supprimé:** suitable
**Supprimé:** of ~15 km x 15 km (= 225 km$^2$)
**Supprimé:** (

Global Climate Models (GCMs) are able to simulate climate conditions at various spatial and temporal scales, whereas climate proxy data are inherently limited by the uneven distribution of sample locations and taphonomic biases (i.e biases in the fossil record, such as pollen preservation, location of archaeological sites, etc.). GCMs use physical equations, e.g. to represent atmospheric fluid dynamics, as well as parameterisations, e.g. for sub-grid scale phenomena, to simulate the Earth's climate. The major disadvantage of GCMs is that they are computationally intensive and are usually only used to model climate behaviour at relatively coarse spatial resolution, typically coarser than 100 km (cf. Flato et al., 2013 for the latest details on the CMIP5 models) especially for long paleoclimatic simulations. Their ability to simulate the small-scale physical processes that drive local surface variables, such as precipitation, is therefore limited (Wood et al., 2004).

Regional Climate Models (RCMs) represent a physically based approach to climate modelling at a finer spatial scale over a specific region of interest (*e.g.*, Liang et al., 2006, Flato et al., 2013). However, RCMs use GCM outputs to set their boundary conditions. They therefore require the explicit modelling of the physical processes at both coarse and fine scales over the whole planet and over region of interest, respectively, and are also computationally demanding. Statistical Downscaling Methods (SDM), on the other hand, are less computationally demanding. SDMs proceed by empirically associating local-scale variables with large-scale atmospheric variables produced by GCMs, and are faster to compute than mechanistic RCMs (Vaittinada Ayar et al., 2015). SDMs fall into four main families: "transfer functions", which directly link large-scale and local-scale variables; "weather typing" methods based on conditioning statistical models on recurrent weather states; "stochastic weather generators" that simulate downscaled values from their (potentially conditional) probability density functions; and "Model Output Statistics" (MOS) methods based on adjusting (i.e., correcting) the statistical distribution of the large-scale GCM simulations in order to generate local-scale variables with the correct statistical properties (e.g., Vaittinada Ayar et al., 2015).

In this study, we explore and refine the capacity of an SDM from the transfer functions family, based on Generalized Additive Modelling (GAM), to compute temperature and precipitation time series at a fine spatial and temporal resolution for the LGM over Western Europe, south of the Fennoscandian ice-sheets. GAM is a non-parametric statistical technique that has proven reliable for capturing non-linear relationships between local- and large-scale variables and correcting the biases specific to a given GCM (e.g., Vrac et al., 2007; Levavasseur et al., 2010). The SDM used here accurately downscales the climatology (i.e., the climate averages over several decades) of temperature and precipitation generated by a GCM for the LGM when calibrated using present-day data (Vrac et al. 2007). Its ability to generate projections of the small-scale temporal patterns necessary to explain the spatial dynamics of prehistoric human populations is untested, however. In the present study, therefore, we use present-day climate data (corresponding to the average of the 1961-1990 period) extracted from the IPSL-CM5A-LR GCM (Dufresne et al., 2013) to calibrate the SDM, applying it to a 50 year-long time series of climate simulations for the LGM (Kageyama et al., 2013a, b). Interpolated values of coarse-grain variables extracted from the GCM, as well as fine-scale geographical data such as elevation and advective continentality, are used as predictors in the GAM. The result is the

**Supprimé:** RCM

**Supprimé:** apply

**Supprimé:** . We use a 50 year-long time series

**Supprimé:** climate simulations for

**Supprimé:** LGM (Kageyama et al., 2013a, b) and the present climate (Dufresne et al., 2013) extracted from the IPSL-CM5A-LR GCM (Dufresne et al., 2013).

**Supprimé:** (e.g., Vrac et al., 2007; Levavasseur et al., 2010). In the present study, interpolated

production of downscaled monthly values over 50 years for temperature and precipitation, including local temporal variability in temperature and precipitation rates. In addition, we compare the impact of three different interpolation techniques (bilinear interpolation, bicubic interpolation and kriging) on the downscaling results, evaluating the resulting SDM outputs with the aid of climate proxies (palynological and archaeozoological data) and observing the impact of each technique on patterns of spatial and temporal variability in temperature and precipitation. The SDM using kriging interpolation is demonstrated to be a good compromise between computational complexity and accuracy, validated on an 11-year present-day time series distinct from the calibration period.

## 2 Materials and Methods

### 2.1 Global Climate Model

The GCM used in this study is the ocean-atmosphere coupled model IPSL-CM5A-LR (Dufresne et al., 2013) developed for the CMIP5 (Taylor et al., 2012) and PMIP3 (Braconnot et al., 2012) projects and the 5th IPCC report (IPCC, 2013). The IPSL-CM5A-LR model has a spatial resolution of 1.9° in latitude and 3.75° in longitude over Europe, which is the area of interest here (i.e. ~62 500 km$^2$). The model performance and main biases are described in Dufresne et al. (2013) and Hourdin et al. (2013). This model version is known to have a cold (-1.4°C) bias in terms of globally averaged temperature and the bias in mean annual temperature over Europe is similar to the global value. In this low-resolution version of the model, the mid-latitude westerly winds are generally more equatorial than observed (Hourdin et al., 2013) and this is the case for the Northeast Atlantic and Europe too. The general pattern of extra-tropical precipitation over the North Atlantic and European sectors is satisfactory with respect to the annual mean (Dufresne et al., 2013). The equilibrium temperature response to a doubling $CO_2$ is 3.59°C (Dufresne et al., 2013), which is rather high compared to other CMIP5 models. Nevertheless, the model's response to LGM forcings somewhat underestimates the cooling over Europe, as reconstructed from pollen, and is satisfactory in terms of precipitation (Kageyama et al., 2013).

We use a *historical* simulation run according to the CMIP5 protocol, and use model output for the period from 1961 to 1990 as our present-day reference climate. Outputs from this simulation are used in the calibration process (below). The simulation of LGM climate conditions follows the PMIP3 protocol (Braconnot et al., 2011, Braconnot et al., 2012). The concentrations of atmospheric greenhouse gases were lowered to their LGM values derived from ice core data (185 ppm for $CO_2$, 350 ppb for $CH_4$ and 200 ppb for $N_2O$) and the ice sheets are prescribed according to the product developed for PMIP3 (Abe-Ouchi et al., 2015). The model is run for several hundred years until the response to the LGM forcing in terms of surface climate variables is stabilised (Kageyama et al., 2013a, b). For this research, we extracted 50 years of monthly mean data (temperature 2m above the surface, precipitation, sea level pressure and relative humidity) from the stabilised part of the simulation. Next, we downscaled the data, calculated their average climatology and their temporal (interannual) variability.

### 2.2 Generalized additive models



Supprimé: . We
Supprimé: the calculation
Supprimé: -
Supprimé: A present-day
Supprimé: was
Supprimé: 1960
Supprimé: with the corresponding forcings in atmospheric composition and land use changes

Generalized additive models (GAM, Hastie and Tibshirani, 1990) are statistical models blending the properties of generalized linear models with additive models. Given a dependent variable $Y$ and $p$ predictor variables $[X_1,\ldots,X_p]$, GAMs compute $E(Y|X_1,\ldots,X_p)$, the expected value of $Y$, conditionally on the $p$ predictors $X_i$, as a sum of non-parametric functions as follows:

$$E(Y|X_1,\ldots,X_p) = \sum_{i=1}^{p} f_i(X_i), \quad (1)$$

Following Vrac et al. (2007), cubic spline functions were used for the $f_i$, represented by piece-wise third-order polynomial functions. For each function $f_i$, a number of knots are placed evenly throughout the predictor range, and the cubic polynomials that compose $f_i$ are constrained to continuity conditions at each knot to ensure smooth transitions (Wood, 2000, 2004). GAMs were calibrated using the mgcv package (Wood, 2011) in R, and the number of knots was determined automatically using generalised cross-validation.

Using a combination of geographical and physical predictor variables has been shown to improve spatial downscaling results (Vrac et al., 2007). The GAM uses these predictor variables, in addition to the original variables to be downscaled, to correct the biases of the GCM described above. This implies that a given GAM is only valid for the GCM for which it was calibrated, as it corrects its' specific biases. Using GAMs on climate variables requires the predictor and dependent variables to have the same spatial scale. The present-day dependent variables (precipitation and temperature) are at a fine spatial scale. The elevation variable is also at a fine spatial scale, whereas the predictor climate variables generated by the GCM are at a coarser scale. Thus, interpolation of the predictor climate variables is necessary. For this research we tested three interpolation techniques (see below).

Once the functions $f_i$ have been fitted using the present-day data, the downscaling can be performed on the GCM outputs for the LGM. The downscaling uses fine-scale and interpolated predictor climate variables corresponding to the LGM to generate fine-scale dependent variables. Here, two GAMs are calibrated: one for temperature and one for precipitation.

### 2.3 Calibration data

#### 2.3.1 Fine-scale dependent variables: the CRU climatology.

Fine-scale, present-day temperature and precipitation dependent data were obtained from the Climate Research Unit (CRU, New et al., 2002). The spatial resolution of the data is 10' (*i.e.* 1/6 degree), regularly gridded between 32.72° and 59.861° latitude (N = 164 values) and -11.578° and 24.738° longitude (N = 219 values) for a total of N = 35916 grid-points. We computed a monthly climatology for each gridpoint by averaging the variable of interest (temperature or precipitation) over 30 years (from 1961 to 1990) for each month (Fig. S1), resulting in 12 values for each cell. The GAM is calibrated over this 30-year climatology, because the GCM cannot be set up to generate predictor variables for a specific year.

**Supprimé:** 2006

**Supprimé:**

**Mis en forme:** Anglais (Canada)

**Supprimé:** 2000

**Supprimé:** 31

**Supprimé:** 1960

**Supprimé:** S1).

**Mis en forme:** Police :Pas Gras, Anglais (G.B.)

We are specifically interested in downscaling temperature and precipitation over Western Europe, south of the ice-sheets during the LGM, a time when archaeological data indicates that human populations contracted in size and range. The GAMs were calibrated over a wider area than the region of interest in order to avoid edge effects and include the full range of climate conditions that prevailed during the LGM, which was much colder than the present day. As a result, the calibration domain extends further North and East (where more continental conditions prevail) than the region of interest. Preliminary simulations nonetheless showed that selecting too large a calibration region averaged-out the small-scale variation we are interested in.

**2.3.2 Large-scale predictor variables.**

We used the data from a CMIP5 historical simulation run with the IPSL-CM5A-LR model to produce the predictor climate variables used for the calibration of the GAMs. We calculated monthly climatological averages from the simulation outputs for the period from 1961 to 1990, i.e. the same years as the CRU data (see above). The predictor variables we used are: temperature (T), precipitation (P), atmospheric pressure at sea level (SLP) and relative humidity (RH). The variables were spatially interpolated to match the spatial resolution of CRU data, which is 10'; each grid-point in the CRU data therefore matches a value for each of the predictor variables. Three interpolation methods were tested: bilinear, bicubic and kriging (Figs. S2-S4). For the kriging method, we used the "krig" function from the vacumm python package (http://relay.actimar.fr/~raynaud/vacumm/) using an exponential fit of the variogram, with the fit computed independently for every month and every variable. Different interpolation methods generate differences in the fine-scale predictor data. For example, the bicubic interpolation generates values outside of the initial range of values, contrary to the other two techniques. The bilinear interpolation generates abrupt changes in the slope of the values at the initial data points, whereas the other two techniques generate smooth surfaces. It is therefore important to assess the potential impact of the interpolation method on the output of the downscaling process.

**2.3.3 Fine-scale predictor variables.**

We extracted present-day elevation data from the CRU dataset gridded at the same fine-scale spatial resolution as the dependent variables. We computed the advective (Aco) and diffusive (Dco) continentalities, following Vrac et al. (2007). Dco is bounded between 0 and 1, and corresponds to the shortest distance to the ocean. A low value means that distance to the ocean is small, and *vice versa*. Aco is also bounded between 0 and 1, and takes the direction and intensity of prevailing winds into account along with the distance to the ocean. The change of Aco during a time *dt* is governed by Equation 2:

$$dAco = [-Aco(1 - i_{co}) + (1 - Aco)i_{co}]\frac{\frac{U}{l_0}}{\frac{U}{U_0}}\ln 2 \; dx \qquad (2)$$

Supprimé: 1960

Supprimé: for

Supprimé: interpolated variable.

where $i_{co}$ is 0 over sea and 1 over land, *dx* is the distance traveled by the air mass during *dt*, *U* is the mean wind norm, obtained from the GCM, and $l_0/U_0$ is the distance/wind ratio corresponding to a change of Aco of 1/2. Both variables are used to account for the fact that an air mass becomes more continental as it travels across land. Since Dco and the Aco proved to be highly correlated but Aco provided the best performance in the models, we only selected Aco for the present analysis (Figs. S2-S4).

**Mis en forme:** Police :Pas Gras

### 2.4 Calibration of the GAMs

For each dependent variable (temperature and precipitation) and for each interpolation technique (bilinear, bicubic and kriging), we tested different combinations of physical and geographical predictor variables. To downscale temperature, we computed GAMs for all possible combinations of coarse-grain temperature values from the GCM interpolated at fine scale, with fine-grain elevation and advective continentality (Aco), resulting in seven possible models for each interpolation. To downscale precipitation, we computed GAMs for all possible combinations of coarse-grain temperature, coarse-grain precipitation, sea-level pressure, and relative humidity values from the GCM interpolated at fine scale, with fine-grain elevation and advective continentality (Aco), resulting in 31 possible models for each interpolation. For each interpolation technique, the resulting GAMs were compared using the Akaike Information Criterion (AIC; Akaike, 1974), and the model with the lowest AIC was selected. The AIC is a measure of the relative goodness of fit of each of the models and penalizes the number of parameters, thus preventing overfitting. The significance of each variable was assessed using p-values, and verified by visual inspection of the spline 95% confidence intervals. Six GAMs were therefore retained after calibration (one for each response variable and for each interpolation technique; Table S1).

**Supprimé:** )

### 2.5 Downscaling of temperature and precipitation time series for the LGM

We computed downscaled temperature and precipitation values using the six GAMs resulting from the calibration process described above, i.e. for the same predictor variables. The large-scale climate variables were generated by the GCM using the PMIP3 protocol for the LGM prior to interpolation (Figs. S5-S7). The geographical variables are derived from a digital elevation model for the LGM (Levavasseur et al., 2011). In particular, the change in coastlines due to the lower sea-level at LGM is accounted for, which has an impact on the continentalities. The downscaling was performed for each month of the 50-year-long monthly output from the GCM, in order to obtain a long time series of fine-scale temperature and precipitation values over Europe and calculate climatological averages. We also calculated indices of variability, including measures of variance and inter-annual variability for the variables of interest. The standard deviation of monthly mean temperatures for each month was calculated for the 50-year run. The coefficient of variation of monthly mean precipitation values was calculated for the same period.

**Déplacé (insertion) [1]**

**Mis en forme:** Anglais (G.B.)

**Supprimé:** as well as variability and extremes

### 2.6 Evaluation data (palynological data and vertebrate remains)

**Mis en forme:** Interligne : 1,5 ligne

**Mis en forme:** Police :Pas Gras

To evaluate the performance of the SDM for the LGM we compared our temperature and precipitation outputs to local climate variables estimated on the basis of pollen and vertebrate fossils from 29 test locations (Table S1, Fig. 1). For each of our 29 test sites, we estimated the mean, minimum, and maximum temperature and precipitation rate on a monthly basis over the course of the 50 downscaled years, and compared the ranges of downscaled values to the ranges of temperature and precipitation values reconstructed using the palynological data and vertebrate remains.

Reconstruction of local temperature and precipitation values (annual mean temperature, mean temperature of the coldest month, mean temperature of the warmest month, mean annual precipitation, precipitation in January, precipitation in July) were obtained from pollen data reported in an independent study using inverse vegetation modelling for 14 sites located in Europe (Wu et al., 2007). For the remaining 19 sites, vertebrate remains from another study (Burke et al. 2014) were used to calculate bio-climate indices (BCI; ff. Hernandez Fernandez, 2001a, b). The method set forth by Hernandez-Fernandez uses large and small vertebrates to compute the relative probability that a given assemblage reflects one of Walter's nine global zonobiomes (Walter and Box, 1976). The method is based on the "climate envelope" method commonly employed in biogeographical reconstructions. The BCI uses presence/absence data, thus avoiding the problems inherent with calculating the relative representation of species from the archaeozoological record, and all available taxa rather than one or two "indicator" species, thus avoiding the risk that changes in the distribution of a single taxon could bias the biogeographical reconstruction. Ranges of temperature and precipitation values (minimum, mean and maximum) for each zonobiome were estimated by mapping the modern distribution of zonobiomes in the northern hemisphere and compiling present-day temperature and precipitation data from the CRU data (see Burke et al. 2014). The zonobiomes (calculated using BCI) were then used to predict the climate ranges for each test location. Note that the range of values for each zonobiome corresponds to the minimum and maximum values for the zonobiome over Western Europe (see Burke et al. 2014). It is therefore not specific to a given location and encompasses a large range of values. The intervals generated by these two different climate reconstruction methods, therefore, are not equivalent. Nevertheless, they provide useful references for evaluating the values generated by the SDM.

## 2.7 Validation using present-day data (CRU 1950-1960 time series)

Due to computational constraints, the validation of the SDM was performed for the kriging interpolation technique only, based on the comparison of the results of the downscaling for the LGM between the three interpolation techniques (see below). A GCM simulation was produced for the period from 1950-1960 and compared with a time series of temperature and precipitation at 10' resolution over Europe for the same period (Mitchell et al. 2004). Yearly averages and variability indices (standard deviation for temperature and coefficient of variation for precipitation) for the two sets of data were compared. The time series was based on the same original data used to create the 1961-1990 climatology, which forms the calibration set for the SDM (New et al. 1999). The time series was created by spatially interpolating data from irregularly spaced climate stations using

Supprimé: ,
Supprimé: S8
Supprimé: , simulated
Supprimé: , simulated
Supprimé: 10
Supprimé: the Iberian peninsula
Supprimé: Because archaeozoological assemblages are thanatocoenoses, rather than biocoenoses, they may represent more than one biome. Thus, we considered the two most probable zonobiomes (the two highest BCIs) for each site
Mis en forme: Barré
Déplacé (insertion) [2]
Supprimé: 2.7 Variability indices

thin-plate smoothing splines, however, and may be subject to its own bias. Potential differences between our results and the validation time series should therefore be interpreted with caution.

## 3. Results

### 3.1 The GAM

The best models (i.e., the models with the lowest AIC value) for temperature and for precipitation were obtained by using the same sets of variables (one for temperature, one for precipitation) in the GAM for the three interpolation techniques. The predictors for temperature are: simulated temperature from the GCM, elevation and advective continentality (explaining 95.80%, 95.29% and 95.63% of the variance for the bilinear, bicubic, and kriging interpolations, respectively). For precipitation, the predictors are: simulated temperature, precipitation, sea-level pressure and relative humidity from the GCM, elevation, and advective continentality (explaining 64.65%, 64.79% and 65.43% of the variance for the bilinear, bicubic, and kriging interpolations) (Table 1). The p-values for all variables for all models were < 0.001.

The splines resulting from the calibration process for the downscaling of temperature values show that fine-scale temperature readings are related to the GCM temperature and to elevation in a linear fashion, and the differences between the three interpolation techniques tested are negligible (Fig. 2). The fine-scale temperature is proportional to the GCM temperature but it is inversely proportional to elevation, which means that the coarse-grain temperatures generated by the GCM are higher than fine grain observations in regions of high elevation. This is expected because temperature generally decreases with increasing altitude and because in the coarse grain GCM, it is the average elevation over the grid box that is considered. Although the model including all three predictor variables produced the lowest AIC, advective continentality has a very limited impact on temperature, as the values of $f$(Aco) remain close to 0. When applied outside the range of values for which they are calibrated, GAMs use a linear extrapolation of the splines. The range of values for elevation is similar for the present-day and the LGM (Fig. 3). Because of the increased land mass during the LGM (which correlates with a low sea-stand), there are more high values for advective continentality, but this difference has a small impact since the spline is relatively flat for this variable. As expected, temperature is lower during the LGM than for present-day. This has a limited impact on the projections because the linear interpolation of the spline outside of the range of values used for calibration is consistent with the fact that the spline is relatively linear for temperature, and the few remaining values are within 10 degrees of the minimum temperature. For very low temperatures during the LGM, however, the SDM outputs should be interpreted carefully, as discussed below.

The splines for the downscaling of precipitation, in contrast, are non-linear (Fig. 4). The splines showing the influence of temperature on expected precipitation rates show larger variations due to a low expected precipitation for both low and high temperatures, but high expected precipitation for middle-range temperatures. Although the expected precipitation increases monotonically with the interpolated precipitation rates, the spline values are lower than the GCM precipitation values and the

**Supprimé:** . Next, we used the Standardized Precipitation Index (Agnew 2000; Guttman 1999; McKee et al. 1993) to calculate how often a given month deviated from expected precipitation values based on the climatic norm (calculated over 50 years) using a subroutine of the "SPEI" package in R (Vicente-Serrano et al. 2010) and a 12-month interval to standardize the values. Standardized Precipitation Index (SPI) values were then classified as "normal", "very" dry/wet, "severe" dry/wet and "extreme" dry/wet following McKee et al. (1993) and the number of "normal" months was summed to produce the variability index for Precipitation. The Standardized Temperature Index (STI), which is based on the same principle as the SPI, was calculated using the "STI" package in R (Fasel 2014) and used to produce a variability index for temperature.

**Mis en forme:** Anglais (G.B.)

**Supprimé:** 1

**Supprimé:** GCM overestimates

**Supprimé:** for

**Supprimé:** elevations

**Mis en forme:** Police :10 pt

**Supprimé:** were

**Supprimé:** periods

**Supprimé:** 2

**Supprimé:** should have

**Supprimé:** should nonetheless have

**Supprimé:** .

**Supprimé:** 3

relationship is non-linear. The expected precipitation increases more rapidly for low than for high interpolated precipitation, in keeping with previous observations that GCMs (and even RCMs) overestimate drizzles, which may explain this correction (e.g. Gutowski et al.,2013). The three interpolation techniques produced similar splines for all variables, although the splines of the bicubic interpolation are slightly distinct from the other two interpolation techniques. The main difference is observed for the spline of the bicubic relative humidity, which indicates lower precipitation for low relative humidity than the other two interpolation techniques. Differences between the splines of the bicubic interpolation and the other two techniques were expected, since this interpolation generates the most divergent values (Figures S8, S11). As a GAM will adjust the splines to compensate for the potential biases of a GCM, it will also do so to compensate for the specificity of an interpolation technique.

Advective continentality is the variable with the least impact on precipitation rates. The ranges of values for simulated precipitation, relative humidity, and for elevation are similar for the present-day and the LGM periods, and the distributions of the variable substantially overlap, indicating that the splines calibrated over the present-day period can apply for the LGM (Fig. 3). As for temperature, the spline of advective continentality is relatively flat, and the difference of range of values will have limited impact on the projections. The spline for the simulated atmospheric pressure at sea level has a positive slope for high values. This spline does not represent a causal relationship, but simply indicates that the GCM tends to underestimate precipitation at high atmospheric pressure. The simulated atmospheric pressure at sea level is also higher for the LGM than for the present-day period. Nonetheless, since the atmospheric pressure is mostly lower than 1045 hPa during the LGM, and given the low slope of the spline on the right-hand extremity, this discrepancy should have little impact on the results. For temperature, the splines are relatively linear near the lower end of the range of present-day values, and the linear interpolation of the spline at lower values for the LGM is therefore sensible.

### 3.2 Results for the LGM

#### 3.2.1 Temperature

Downscaled annual mean temperature was very similar for the three interpolation techniques tested (Fig. 5). This was expected, since the splines for the GCM temperature and elevation for all three techniques are also very similar. Temperatures interpolated with the bilinear and kriging techniques were more similar to each other than to the temperature interpolated using the bicubic technique before (Fig. S8) and after (Fig. 5) applying the GAMs. The differences between the bilinear and kriging techniques show a pattern corresponding to the original coarse-grain cells from the GCM. This illustrates the difference between the two interpolation techniques: kriging generates smoother variations than the bilinear interpolation, which generates discontinuous variations at the original points. This difference remained after applying the GAMs, showing the impact of the interpolation technique used on the final outcome of the downscaling.

Supprimé: 2
Supprimé: Similarly,
Supprimé: was lower
Supprimé: , but the corresponding spline calibrated on present-day data is linear at low values
Supprimé: has
Supprimé: influence
Supprimé: projections
Supprimé: Projections
Supprimé: Figs. 4,
Supprimé: Figs. S9, S10
Supprimé: were regularly spaced,

The main differences between the interpolated and downscaled temperatures occur in the northeast of Europe (Fig. S9), where downscaled temperatures are higher, especially in winter. This difference was also observed when comparing present-day CRU data with the interpolated GCM data (Fig. S10), although with a much lower amplitude. Northeast Europe is the coldest region of the study area for both the LGM and the present-day (Figs. S1-S7). Since the spline for temperature has a slope lower than 1 for low temperatures (Fig. 2), the GAM generates higher temperatures than the interpolated values, especially for very low temperatures which fall outside of the present-day range of values due to the linear interpolation of the spline. As a result, the difference between interpolated and downscaled temperatures during the LGM is lower in summer. The SDM also takes fine scale variations in topography into account, such as abrupt elevation changes in the Alps and Pyrenees (Fig. S9).

The range of temperatures for the 19 sites for which the BCIs were computed are in accordance with the temperature reconstructions, irrespective of the interpolation technique used (Fig. 6). Simulated temperature ranges fall within the reconstructed ranges corresponding to the BCIs and are within the reconstructed ranges from Wu et al. (2007) for all test sites, as shown by the overlap of the red and blue error bars with the diagonal (Fig. 6a,c,e). As noted above, since the BCI reconstruction considers the minimum, mean and maximum values for each zonobiome over the whole of Western Europe (see Burke et al. 2014), some downscaled temperature values may differ from the mean, but as long as the error bars overlap with the diagonal are still in accordance with the BCI climate reconstruction.

### 3.2.2 Precipitation

The three interpolation techniques tested here produce similar distributions of precipitation rates but, compared to the two other interpolation techniques, the bicubic interpolation produced the most divergent results (Fig. S11). All three interpolation techniques nonetheless reflect the biases of the GCM (Fig. S12). Precipitation rates predicted using bicubic interpolation were substantially lower than those predicted using the other techniques in the South-West of Europe in winter and higher in the North of Europe (over the current North Sea) in summer (Fig. 7). This result is consistent with the observation that the splines for the bicubic interpolation differed from the other two interpolation techniques. Despite a general agreement between simulated and observed annual precipitation mean over Europe (Dufresne et al. 2013), comparing GCM projections interpolated at fine scale with the present-day CRU data shows that the GCM overestimates precipitation over the South of Europe in winter and underestimates them over the North in summer (Fig. S13), which results in the non-linear spline for precipitation (Fig. 4). Coarse-grain GCMs are known to perform poorly when simulating the small-scale physical processes that drive local surface variables such as precipitation (Wood et al., 2004). This explains the discrepancies between the present-day simulations and the CRU data and, by extension, explains the adjustments performed by the SDM.

The precipitation ranges for the 19 sites for which the BCIs were computed are in accordance with the precipitation ranges reconstructed for the LGM. Simulated precipitation ranges for all sites fall within the reconstructed ranges corresponding to the BCIs, as shown by the overlap of the horizontal error bars (in red) with the diagonal (Fig. 8). The SDM predicts higher

**Supprimé:** North-East region
**Supprimé:** S11), with the
**Supprimé:** temperature being
**Supprimé:** S12
**Supprimé:** North-East
**Supprimé:** S2
**Supprimé:** 1
**Supprimé:** falling
**Supprimé:** the
**Supprimé:** temperature
**Supprimé:** took
**Supprimé:** S11
**Supprimé:** primary BCIs for all sites (Fig. 6a,c,e). The reconstructions from Wu et al. (2007) produce smaller ranges of values, but the simulated ranges overlap with reconstructed temperatures (Fig. 6b,d,f).
**Supprimé:** Bilinear
**Supprimé:** and kriging resulted in very
**Supprimé:** downscaled precipitation values, whereas the
**Supprimé:** Summer (Figs
**Supprimé:** , 8
**Supprimé:** techniques. The three
**Supprimé:** produced similar distributions of
**Supprimé:** rates, but compared to the two other interpolation techniques the results of the bicubic interpolation differed the most (Figs. S13, S14). A comparison of interpolated and downscaled precipitation for the LGM with all three interpolation techniques shows that the GCM underestimates precipitation over the South of Europe and overestimates it over the North in winter, and substantially underestimate precipitation over all of Europe in Summer (Fig. S15). Comparing
**Supprimé:** confirms this pattern in winter, and
**Supprimé:** Summer
**Supprimé:** S16
**Supprimé:** 3). This is due to the fact that
**Déplacé vers le haut [2]:** 2004).
**Supprimé:**
**Supprimé:** Precipitation
**Supprimé:** primary and secondary BCIs (Fig. 9). As with temperature, the reconstructions from Wu et al. (2007) show smaller ranges of precipitation than the reconstructions using the BCIs

precipitation values for 1 site (in North-West Iberia) than the reconstructions provided by Wu et al. (2007) (the simulated mean precipitation of the driest month was predicted to be higher than the maximum precipitation found by Wu et al.). However, while the BCI ranges correspond to minimum and maximum values over a relatively large spatial extent, the reconstructions offered by Wu et al. (2007) are site-specific and therefore produce smaller ranges of values. Moreover, Wu et al. (2007) based their reconstructions on local adjustments of the biome estimates from the BIOME4 model (Kaplan et al. 2003). They therefore used the same initial values for different sites, which may underestimate differences between sites and explain the lower range of precipitation values compared to the values generated by the SDM.

### 3.2.3 Variability

The temporal variability of temperature and precipitation rates highlights differences between the three interpolation techniques. When temporal variation was computed over the interpolated variables (Figs. S14, S15), bilinear interpolation displays regular spatial patterns for both temperature and precipitation, especially in summer. This pattern was less apparent for kriging, and almost absent for bicubic interpolation.

With bilinear interpolation, the interpolated values will necessarily be less variable (whatever the index used) than the original values. For a simple linear interpolation, given 2 spatially consecutive values at 2 different points in time ($y(x_0,t_0)$, $y(x_0,t_1)$, $y(x_1,t_0)$ and $y(x_1,t_1)$), any linearly interpolated value $y(x_i)$ for a location $x_i$ in $[x_0, x_1]$ will necessarily be comprised in $[y(x_0),y(x_1)]$. In addition, due to the linear interpolation $|y(x_i,t_1)- y(x_i,t_0)| < |y(x_0,t_1)- y(x_0,t_0)|$ and $|y(x_i,t_1)- y(x_i,t_0)| < |y(x_1,t_1)- y(x_1,t_0)|$. By contrast, since kriging does not impose linear interpolation between $x_0$ and $x_1$, this relation does not necessarily hold, and even less for bicubic interpolation, which do not impose $y(x_0) \leq y(x_1) \leq y(x_1)$. However, because of this absence of restriction, bicubic interpolation can generate values with a high variability, as for precipitation in summer in the South-West of the Iberian peninsula (Fig. S15; the high variability for kriging in winter only occurs at the boundary of the study area due to boundary conditions, and can therefore be discarded).

Computing temporal variation over the downscaled variables (Figs. 9, 10), showed that the GAMs attenuated this spatial pattern, which nonetheless still occurred for the bilinear interpolation, and was almost non-existent for the other two interpolation techniques.

### 3.3 Validation on present-day data

The comparison of average downscaled (based on kriging interpolation) and observed (CRU) monthly temperature and daily precipitation for the 1950-1960 period shows that they are in good agreement (Figures S16, S17). For temperature, the main difference occurs in the far North of the study area (Southern point of the Scandinavian peninsula), and on the Italian side of the Alps. The downscaled temperature is similar on both sides of the Alps, whereas there is some difference in the CRU data

**Supprimé:** showed important

**Supprimé:** S17, S18

**Supprimé:** showed

**Supprimé:** have a lower variation

**Mis en forme:** Français

**Supprimé:** S18

**Supprimé:** , 11

**Supprimé:** patterns

**Supprimé:** . ... [1]

(Figure S16), suggesting the inclusion of orographic wind as a predictor may improve the SDM for specific areas. For precipitation, the main difference occurs in areas of high precipitation, especially the Western coast of Great Britain (Figure S17). Nonetheless, these areas were the areas with higher levels of precipitation for both datasets.

The overall patterns of variability were overall similar between the downscaled and CRU datasets (Figures S18, S19), with nonetheless some local differences. Temperature variability was higher in the North-East region of the study area and lower in the South-West region, especially in winter (Figure S18), and the range of values for the standard deviation were very similar for the different seasons. Downscaling tended to slightly overestimate temperature variability in the North of the study area, and underestimate it in the North-East compared to the CRU data. Some small-scale differences are nonetheless difficult to interpret, since the CRU data showed some spatial artifact, for example in the center of the Iberia peninsula. The amplitude of the variability values was more different for precipitation, with the variability observed in the CRU data being higher (Figure S19). Nonetheless, the general patterns were quite similar, with the Southern region of the study area having higher precipitation variability than the North for both the downscaled and the CRU datasets.

**4 Conclusion and Discussion**

We downscaled temperature and precipitation values produced by the IPSL-CM5 model for 50 simulated years over Western Europe during the last glacial maximum using a GAM, a computationally efficient method for downscaling GCMs (Vrac et al. 2007). A single GAM was used for each dependent variable, calibrated over an average of 30 years of present-day data. Comparing the outputs of the SDM with two different climate reconstructions showed that this method generates results that fall within the computed confidence intervals for the variables of interest. This enabled us to compute indices of climate variability for the LGM in Western Europe. In a separate study, we were then able to test a suite of environmental predictors and demonstrate that climate variability is a key factor governing the spatial distribution of prehistoric human populations during the LGM (Burke et al. 2014, 2017).

Downscaled time series for a present-day period (1950-1960) falling outside of the calibration period (1961-1990) were in good agreement with an independent time series for both averaged values and measures of variability. Our study, therefore, demonstrates that the SDM, originally designed to downscale climatology data (averaged over several decades), can be applied to a time series thus allowing us to compute spatio-temporal patterns at a fine scale appropriate for studying the spatial dynamics of prehistoric human populations. SDMs must be carefully parameterised, however, including selecting the appropriate size of the area used for calibration.

Overall, the downscaled temperature and precipitation values produced by the SDM are in agreement with the climate reconstructions obtained from vertebrate remains and palynological data, with few exceptions (Figures 6, 8). The SDM results for the LGM differ from the interpolated data in the northeast of the study area, reflecting the adjustments made in the GAM

**Supprimé:** We apply GAM-based statistical downscaling, a computationally efficient method for the downscaling of large-scale climatic variables from global climatic models (Vrac et al. 2007), to the downscaling of

**Supprimé:** .

**Supprimé:** the

**Supprimé:** 31

**Supprimé:** the

**Supprimé:** produced satisfying

**Supprimé:** . It also

**Supprimé:** . Elsewhere, we

**Supprimé:** society

**Supprimé:** Nonetheless, SDMs require a careful evaluation of the different variables considered.

**Supprimé:** The

to counter biases inherent in the IPSL-CM5A-LR GCM used in this study. These discrepancies have little consequence in the present study, since this region was covered by ice-sheets during the LGM. It was included for calibration because it represented present-day climate conditions that were close to those of the Southern part of the study area during the LGM. This region was also included in the downscaling to illustrate the fact that, since it was colder during the LGM than any present-day region of the study area, results for this region should be interpreted with caution.

The choice of interpolation technique used in the SDM also proved critical as it has a strong impact on the distribution of climate variability. We tested three different interpolation techniques. Since GAMs require the predictor and the dependent variables to have the same spatial grain, bilinear interpolation is commonly used to downscale the coarse-grain data generated by GCMs (Vrac 2007). However, as this research demonstrates, bilinear interpolation generates non-smooth surfaces which may cause spatial artifacts in the final output. We tested two other non-linear interpolation techniques which generate smoother surfaces: bicubic interpolation and kriging. Bicubic interpolation generates values outside of the initial range of values (and therefore under- or overestimates the values) but is faster to apply than kriging. Kriging is more computationally demanding but offers the advantage of constraining the interpolated values within the range of initial values. The three interpolation techniques produced different results for both temperature and precipitation during the LGM (Figs. S8, S11). After applying the GAM, these differences were especially important for precipitation values (Fig. 7). Because the GCM generated coarse grain temperature values for present-day conditions which are highly correlated with the CRU data, all three interpolation techniques produced similar linear splines and led to similar results for this variable. In the case of precipitation, however, bicubic interpolation predicts drier environments than the other two techniques by up to 2 mm/day. Since GCMs operate at grains that are too coarse to accurately model small-scale physical processes driving local surface variables (Wood et al., 2004), the SDM for precipitation relies on more variables than are required to model temperature. The splines for these variables are non-linear (Figure 4), however, and may exacerbate the differences between the bicubic interpolation and the other two techniques. The cumulative impact of the interpolation and the GAM can therefore be non-negligible. This highlights the utility of the comparison presented in this research, especially for local phenomena such as precipitation.

The variability maps produced when using bilinear interpolation show the presence of a spatial artefact, in the form of a regular grid, for both temperature and precipitation (Figs. S14, S15). This artefact reflects the fact that bilinear interpolation generates lower variability between the points from which the interpolation is performed. Although slightly attenuated, this artefact remained after applying the GAMs (Figs. 9, 10). Prior to applying the GAMs, bicubic interpolation produced maps with the smallest level of artifacts, kriging was intermediate and bilinear interpolation produced the highest level of artifacts. However, bicubic interpolation sometimes generated unrealistically high variability for precipitation (Fig. S15) while the artifacts generated by kriging decreased after applying the GAMs (Figs. 9, 10). We conclude that although more computationally demanding than the other two techniques, kriging represents a good compromise between computational complexity and accuracy. Contrary to bicubic interpolation, kriging generates values within the range of the values generated by the GCM and

**Supprimé:** for the accuracy of the output.
**Supprimé:** the
**Supprimé:** the GCM
**Supprimé:** is shown to generate
**Supprimé:** ,
**Supprimé:** Kriging
**Supprimé:** showed differences
**Supprimé:** S10, S14
**Supprimé:** 8). In particular
**Supprimé:** the
**Supprimé:** , which
**Supprimé:** explain
**Supprimé:** important
**Supprimé:** SDM based on
**Supprimé:** SDM based on the
**Supprimé:** pinpoints
**Supprimé:** here
**Supprimé:** Moreover, the
**Supprimé:** S17, S18
**Supprimé:** , 11
**Supprimé:** and
**Supprimé:** S18). After applying the GAMs, the artifact
**Supprimé:** , 11). .
Because the GCM generated reliable temperatures at coarse grain which were highly correlated with the CRU present-day temperatures, the three interpolation techniques produced similar linear splines and led to relatively similar values for this variable.
**Supprimé:** seems to represent

generates variability indices with more realistic patterns than the bilinear interpolation. We therefore recommend using kriging for SDM applications based on the method presented here.

The IPSL-CM5A-LR GCM is known to predict lower temperatures than the values observed at high latitudes in winter (Dufresne et al., 2013). This bias was indeed observed when comparing the interpolated temperature with the CRU present-day data. As a result, the spline for temperature has a shallow slope at low temperatures (Fig. 2). The resulting correction applied by the GAM was emphasised for the LGM data generated by the GCM in winter in the North of Europe (Fig. S9), which lie outside of the range of present-day temperature and therefore relied on a linear interpolation of the spline. The large differences in temperature are therefore likely to be a combination of an underestimation of temperature by the GCM, and an over-correction of the very low temperature by the SDM. The spatial domain used to calibrate the GAM is larger than the domain of interest, namely Western Europe south of the ice sheets (the region occupied by human populations during the LGM) for reasons discussed above. These include the necessity of avoiding edge effects and including the full range of climate conditions likely to have occurred during the LGM. The observed over-correction lies on the periphery of the calibration region and is not within the study region this SDM was designed for.

The calibration area selected should therefore be large enough to encompass a representative range of climate conditions but should overlap the study region in order to account for potential relationships between climate and geographical variables specific to the region. However, through trial and error we found that using too large a calibration region averages out these relationships and therefore runs counter to the objectives of the downscaling, which is to represent fine scale spatial variation. Further research into the impact of the size of the calibration region on the SDM would be an interesting avenue to pursue. In addition, some results presented here are probably highly influenced by the specific GCM that was used. Especially, the variability in the SDM results is sontrgly influenced by the variability of the original GCM, in addition to the choice of the interpolation technique (Figure 9, S14, 10, S15). For temperature, for example, for a given interpolation technique, the downscaling adjusts the interpolated GCM temperature based on elevation, which is constant for a given location. The SDM will therefore not change variability compared with interpolation of the GCM. Using ensembles of models increases confidence in climate projections by enabling a better quantification of such uncertainty (Tebaldi & Knutti 2007). Although the outputs of ensembles of models may be challenging to interpret, this is another promising avenue for improving the application of the SDM method presented here that should be pursued in the future, especially for the computation of variability indices.

Our goal in this research has been to develop and test tools for the production of climate simulations at suitable spatial and temporal scales for investigating the mechanisms through which climate change and climate variability may have affected human populations in the past. Our aim is to help explain some of the broad evolutionary patterns visible in the archaeological record. Our results demonstrate the potential of GAMs for the production of climate simulations at a fine scale of resolution, both spatially and temporally, at relatively low computational cost. The resulting climate simulations can be used to test human

**Supprimé:** more reliable

**Supprimé:** Because the GCM generated reliable temperatures at coarse grain which were highly correlated with the CRU present-day temperatures, the three interpolation techniques produced similar linear splines and led to relatively similar values for this variable.

**Supprimé:** had

**Supprimé:** temperature

**Supprimé:** 1). This

**Supprimé:** S11

**Supprimé:** are

**Supprimé:** ,

**Supprimé:** No palynological or vertebrate data were available

**Supprimé:** evaluate the performance

**Supprimé:** downscaling method in this region

**Supprimé:** such a

**Supprimé:** should be treated with caution. For

**Supprimé:** purpose of studying

**Supprimé:** distribution of modern human population, this overcorrection will have negligible effects, since this region was covered by an ice cap during

**Supprimé:** time

**Supprimé:** and the range

**Supprimé:** values over

**Supprimé:** whole region in

**Supprimé:** present-day data encompasses the range of values

**Supprimé:** region where humans were present

**Supprimé:** (Figs. S2-S7).

**Supprimé:** a

**Supprimé:** scale to investigate

**Supprimé:** locally decision-making

**Supprimé:** broader

**Supprimé:** patterning

decision-making at regional and local scales useful for investigating the spatial distribution of prehistoric populations against a backdrop of inter and intra-annual climate variability (e.g., Burke et al. 2017; Burke et al. 2014).

*Code and data availability.* The code used for the downscaling and the input and output data are available at at https://figshare.com/s/1b952e47ff274cc0687e (DOI: 10.6084/m9.figshare.5487145).

*Author contributions.* GLa, AB, MK, MV and GLe conceived the study. GLa and GLe implemented the R code for the downscaling. MK, MV and GR ran the IPSL-CM5A-LR simulations and implemented the code for the interpolations. CD implemented the code for the computation of the continentality. GL wrote the paper and all commented on it.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* This research was supported by the *Fonds Québécois de Recherche Société et Culture - Programme de soutien aux équipes* (demande no.179537). Jérôme Servonnat is thanked for his input on the IPSL-CM5A-LR model biases.

**References**

Abe-Ouchi, A., Saito, F., Kageyama, M., Braconnot, P., Harrison, S. P., Lambeck, K., Otto-Bliesner, B. L., Peltier, W. R., Tarasov, L., Peterschmitt, J.-Y., and Takahashi, K. Ice-sheet configuration in the CMIP5/PMIP3 Last Glacial Maximum experiments, Geosci. Model Dev., 8, 3621-3637 2015.

Akaike, H. A new look at the statistical model identification. IEEE T. Automat. Contr., 19(6), 716-723, 1974.

Banks, W.E., d'Errico, F., Peterson, A.T., Kageyama, M., Sima, A., and Sánchez-Goñi, M.F. Neanderthal extinction by competitive exclusion. PloS ONE, 3, e3972, 2008a.

Banks, W.E., d'Errico, F., Peterson, A.T., Vanhaeren, M., Kageyama, M., Sepulchre, P., Ramstein, G., Jost, A., and Lunt, D. Human ecological niches and ranges during the LGM in Europe derived from an application of eco-cultural niche modeling. J. Archaeol. Sci., 35, 481-491, 2008b.

Benito, BM., Svenning, J.-C., Kellberg-Nielsen, T., Riede, F., Gil-Romera, G., Mailund, T., Kjaergaard, P.C., and Sandel, B.S. The ecological niche and distribution of Neanderthals during the Last Interglacial. J. Biogeogr., 44(1):51-61, 2017.



**Supprimé:** , particularly with regards to

**Supprimé:** Burke et al. 2017; Burke et al. 2014). The impact of climate change on human decision-making processes at a local scale, specifically the selection of seasonal ranges within annual territories, will have had a cumulative effect leading to the larger-scale patterning currently discussed in the literature. Paleoclimatological models, such as the ones produced here, also allow climatologists to investigate climate events that are not observable in the present – such as sudden climate change – but which may occur in the near future. Collaboration with archaeologists, who produce chronologically controlled datasets, allowed us to evaluate the simulation outputs at a fine scale of resolution.

**Mis en forme:** Anglais (Canada)

**Supprimé:** Agnew, CT. Using the SPI to Identify Drought. Drought Network News (1994-2001), 5, 2000.

Braconnot, P., Harrison, S. P., Otto-Bliesner, B., Abe-Ouchi, A., Jungclaus, J., and Peterschmitt, J.-Y. The Paleoclimate Modeling Intercomparison Project contribution to CMIP5, CLIVAR Exchanges No. 56, Vol. 16, No.2, May 2011, pp 15-19, 2011

Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Masson-Delmotte, V., Abe-Ouchi, A., Otto-Bliesner, B., and Zhao, Y. Evaluation of climate models using palaeoclimatic data, Nat. Clim. Change, 2(6), 417-424, 2012.

Bradtmöller, M., Pastoors, A., Weninger, B., and Weniger, G. C. The repeated replacement model–rapid climate change and population dynamics in Late Pleistocene Europe. Quatern. Int., 247, 38-49, 2012.

Burke, A., Levavasseur, G., James, P.M.A., Guiducci, D., Izquierdo, M., Bourgeon, L., Ramstein, G., and Vrac, M. Exploring the impact of climate variability during the Last Glacial Maximum on the pattern of human occupation of Iberia. J. Hum. Evol., 73:35-46, 2014.

Burke, A., Kageyama, M., Latombe, G., Fasel, M., Vrac, M., Ramstein, G., and James, P.M.A. Risky business: The impact of climate and climate variability on human population dynamics in Western Europe during the Last Glacial Maximum. Quaternary Sci. Rev., 164:217-229, 2017.

Davies, W. and Gollop, P. The human presence in europe during the Last Glacial Period II. Climate tolerance and climate preferences of Mid-and Late Glacial hominids. Neanderthals and modern humans in the European landscape of the last glaciation, 131-146, 2003.

Delgado Huertas, A., Iacumin, P. and Longinelli, A. A stable isotope study of fossil mammal remains from the Paglicci cave, southern Italy, 13 to 33 ka BP: palaeoclimatological considerations. Chem. Geol., 141(3–4), p. 211-223, 1997.

Dufresne, J.-L., Foujols, M.A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J.-P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J.-Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M.-P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, F. G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., and Viovy, N.. Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. Clim. Dynam., 40, 2123-2165, 2013.

Finlayson, C., Giles Pacheco, F., Rodriguez-Vidal, J., Fa, D.A., Maria Gutierrez Lopez, J., Santiago Perez, A., Finlayson, G., Allue, E., Baena Preysler, J., Caceres, I., Carrion, J.S., Fernandez Jalvo, Y., Gleed-Owen, C.P., Jimenez Espejo, F.J., Lopez, P., Antonio Lopez Saez, J., Antonio Riquelme Cantal, J., Sanchez Marco, A., Giles Guzman, F., Brown, K., Fuentes, N., Valarino, C.A., Villalpando, A., Stringer, C.B., Martinez Ruiz, F., and Sakamoto, T. Late survival of Neanderthals at the southernmost extreme of Europe. Nature 443, 850-853, 2006.

Finlayson, C., Fa, D.A., Jiménez Espejo, F., Carrión, J.S., Finlayson, G., Giles Pacheco, F., Rodríguez Vidal, J., Stringer, C., and Martínez Ruiz, F. Gorham's Cave, Gibraltar--The persistence of a Neanderthal population. Quatern. Int., 181, 64-71, 2008.

**Supprimé:** Fasel M. STI: Calculation of the Standardized Temperature Index. R Package version 01 https://cran.r-project.org/web/packages/STI/, 2014.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S.C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M. Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Gamble, C., Davies, W., Pettitt, P., and Richards, M. Climate change and evolving human diversity in Europe during the last glacial. Philos. T. Roy. Soc. B,359(1442), 243-254, 2004.

Gutowski Jr, W. J., Decker, S. G., Donavon, R. A., Pan, Z., Arritt, R. W., and Takle, E. S. Temporal–spatial scales of observed and simulated precipitation in central US climate. J. Climate, 16(22), 3841-3847, 2003.

Guttman, NB. Accepting the Standandardized Precipitation Index: a calculation algorithm. J. Amer. Water Resour. Assoc. 35(2), 311-322, 1999.

Hastie, T. J. and Tibshirani, R. J., 1990. Generalized Additive Models, Chapman and Hall.

Hernandez Fernandez, M. Análisis paleoecológico y paleoclimático de las sucesiones de mamíferos del Plio-Pleistoceno ibérico, Biologie. Universidad Complutense de Madrid, Madrid, p. 368, 2001a.

Hernandez Fernandez, M. Bioclimatic discriminant capacity of terrestrial mammal faunas. Global Ecol. Biogeogr., 10, 189-204, 2001b.

Hourdin, F., Foujols, M. A., Codron, F., Guemas, V., Dufresne, J. L., Bony, S., Denvil, S., Guez, L., Lott, F., Ghattas, J., Braconnot, P., Marti O., Meurdesoif, Y. And Bopp, L. Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model. Climate Dynamics, 40(9-10), 2167-2192, 2013.

Hughes, J.K., Haywood, A., Mithen, S.J., Sellwood, B.W., and Valdes, P.J. Investigating early hominin dispersal patterns: developing a framework for climate data integration. J. Hum. Evol., 53(5):465-474,. 2007.

Huntley, B., Alfano, M. J., Allen, J. R., Pollard, D., Tzedakis, P. C., de Beaulieu, J. L., Grüger, E., and Watts, B.. European vegetation during marine oxygen isotope stage-3. Quaternary Res., 59(2), 195-212, 2003.

Jacumin, P. and Longinelli, A. Relationship between δ18O values for skeletal apatite from reindeer and foxes and yearly mean δ18O values of environmental water. Earth Planet. Sc. Lett., 201(1), p. 213-219, 2002.

IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp, 2013.

Jennings, R., Finlayson, C., Fa, D., and Finlayson, G. Southern Iberia as a refuge for the last Neanderthal populations. J. Biogeogr., 38(10), 1873-1885, 2011.

**Mis en forme:** Français

**Déplacé vers le bas [3]:** Jennings, R., Finlayson, C., Fa, D., and Finlayson, G. Southern Iberia as a refuge for the last Neanderthal populations. J. Biogeogr., 38(10), 1873-1885, 2011. ... [2]

**Déplacé (insertion) [3]**

Jiménez-Espejo, F. J., Martínez-Ruiz, F., Finlayson, C., Paytan, A., Sakamoto, T., Ortega-Huertas, M., Finlayson, C., Iijima, K., Gallego-Torres, D., and Fa, D. Climate forcing and Neanderthal extinction in Southern Iberia: insights from a multiproxy marine record. Quaternary Sci. Rev., 26(7), 836-852, 2007.

Kageyama, M., Braconnot, P., Bopp, L., Caubel, A., Foujols, M.-A., Guilyardi, E., Khodri, M., Lloyd, J., Lombard, F., Mariotti, V., Marti, O., Roy, T., and Woillez, M.-N. Mid-Holocene and Last Glacial Maximum climate simulations with the IPSL model. Part I: comparing IPSL_CM5A to IPSL_CM4. Clim. Dynam., 40, 2447-2468, 2013a.

Kageyama, M., Braconnot, P., Bopp, L., Mariotti, V., Roy, T., Woillez, M.-N., Caubel, A., Foujols, M.-A., Guilyardi, E., Khodri, M., Lloyd, J., Lombard, F., and Marti, O. Mid-Holocene and Last Glacial Maximum climate simulations with the IPSL model. Part II: model-data comparisons. Clim. Dynam., 40, 2469-2495, 2013b.

Kaplan, J. O., Bigelow, N. H., Prentice, I. C., Harrison, S. P., Bartlein, P. J., Christensen, T. R., Cramer, W., Matveyeva, N. V., McGuire, A. D., Murray, D. F., Razzhivin, V. Y., Smith, B., Walker, D. A., Anderson, P. M., Andreev, A. A., Brubaker, L. B., Edwards, M. E., and Lozhkin, A. V. Climate change and arctic ecosystems II: Modeling, paleodata-model comparisons, and future projections. J. Geophys. Res., 108(D19), 8171, 2003.

Levavasseur, G., Vrac, M., Roche, D. M., Paillard, D., Martin, A., and Vandenberghe, J. Present and LGM permafrost from climate simulations: contribution of statistical downscaling, Clim. Past,7, 1225–1246, 2011.

Liang, X.-Z., Pan, J., Zhu, J., Kunkel, K. E., Wang, J. X. L., and Dai, A. Regional climate model downscaling of the u.s. summer climate and future change, J. Geophys. Res.: Atmospheres (1984–2012),111(D10) , 2006.

Longinelli, A., Iacumin, P., Davanzo, S., and Nikolaev, V. Modern reindeer and mice: revised phosphate–water isotope equations. Earth Planet. Sc. Lett., 214(3–4), p. 491-498, 2003.

López-García, J. M., Blain, H. A., Cuenca-Bescós, G., Ruiz-Zapata, M. B., Dorado-Valiño, M., Gil-García, M. J., Valdeolmillos, A., Ortega, A.I., Carretero, J.M., Arsuaga, J.L., Bermúdezde Castroe, J.M., and Carbonell, E. Palaeoenvironmental and palaeoclimatic reconstruction of the latest Pleistocene of El Portalón site, Sierra de Atapuerca, northwestern Spain. Palaeogeogr. Palaeocl., 292(3), 453-464, 2010.

López-García, J. M., Blain, H.-A., Bennàsar, M., Euba, I., Bañuls, S., Bischoff, J., López-Ortega, E., Saladié, P., Uzquiano, P. and Vallverdú, J. A multiproxy reconstruction of the palaeoenvironment and palaeoclimate of the Late Pleistocene in northeastern Iberia: Cova dels Xaragalls, Vimbodí-Poblet, Paratge Natural de Poblet, Catalonia. Boreas, 41(2), 235-249, 2011a.

López-García, J. M., Cuenca-Bescós, G., Blain, H. A., Álvarez-Lao, D., Uzquiano, P., Adán, G., Arbizu, M. and Arsuaga, J. L. Palaeoenvironment and palaeoclimate of the Mousterian–Aurignacian transition in northern Iberia: The small-vertebrate assemblage from Cueva del Conde (Santo Adriano, Asturias). J. Hum. Evol., 61(1), 108-116, 2011b.

López-García, J. M., Cuenca-Bescós, G., Finlayson, C., Brown, K., and Pacheco, F. G. Palaeoenvironmental and palaeoclimatic proxies of the Gorham's cave small mammal sequence, Gibraltar, southern Iberia. Quatern. Int., 243(1), 137-142, 2011c.

McKee, T.D., Doesken, N.J., and Kleist, J. The relationship of drought frequency and duration to time scales. . Proceedings of the 8th Conference of Applied Climatology. Anaheim, California: American Meterological Society. p 179-184, 1993.

Mitchell, T. D., Carter, T. R., Jones, P. D., Hulme, M., & New, M. (2004). A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100). *Tyndall centre for climate change research working paper*, *55*(0), 25.

Mix, A. C., Bard, E., and Schneider, R. Environmental processes of the ice age: land, oceans, glaciers (EPILOG), Quaternary Sci. Rev., 20, 627-657, 2001.

Müller, U.C., Pross, J., Tzedakis, P.C., Gamble, C., Kotthoff, U., Schmiedl, G., Wulf, S., and Christanis, K. The role of climate in the spread of modern humans into Europe. Quaternary Sci. Rev., 30, 273-279, 2011.

New M, Hulme M, Jones P (1999) Representing twentieth- century space-time climate variability. Part I: Develop- ment of a 1961–90 mean monthly terrestrial climatology. J Climate 12:829–856

New, M., Lister, D., Hulme, M., and Makin, I. A high-resolution data set of surface climate over global land areas, Clim. Res., 21(1), 1-25, 2002.

Palombo, M. R., Filippi, M. L., Iacumin, P., Longinelli, A., Barbieri, M., and Maras, A. Coupling tooth microwear and stable isotope analyses for palaeodiet reconstruction: the case study of Late Middle Pleistocene Elephas (Palaeoloxodon) antiquus teeth from Central Italy (Rome area). Quatern. Int., 126–128(0), p. 153-170, 2005.

Sanchez Goñi, M. F., and Harrison, S. P. Millennial-scale climate variability and vegetation changes during the Last Glacial: Concepts and terminology.Quaternary Sci. Rev., 29(21), 2823-2827, 2010.

Schmidt, I., Bradtmöller, M., Kehl, M., Pastoors, A., Tafelmaier, Y., Weninger, B., and Weniger, G. C. Rapid climate change and variability of settlement patterns in Iberia during the Late Pleistocene. Quatern. Int., 274, 179-204, 2012.

Sepulchre, P., Ramstein, G., Kageyama, M., Vanhaeren, M., Krinner, G., Sánchez-Goñi, M. F., and d'Errico, F.. H4 abrupt event and late Neanderthal presence in Iberia. Earth Planet. Sc. Lett., 258(1), 283-292, 2007.

Serangeli, J. and Bolus, M. Out of Europe - The dispersal of a successful European hominin form. Quartar 55, 83-98, 2008.

Shea, J.J. Transitions or turnovers? Climatically-forced extinctions of Homo sapiens and Neanderthals in the east Mediterranean Levant. Quaternary Sci. Rev., 27, 2253-2270, 2008.

Tallavaara, M., Luoto, M., Korhonen, N., Järvinen, H., and Seppä, H. Human population dynamics in Europe over the Last Glacial Maximum. P. Natl. Acad. Sci. USA, 112(27):8232-8237, 2015.

Taylor, K.E., R.J. Stouffer, R.J. and Meehl, G.A. An Overview of CMIP5 and the experiment design." Bull. Amer. Meteor. Soc, 93, 485-498, 2012.

Tebaldi, C., and Knutti, R. The use of the multi-model ensemble in probabilistic climate projections. Philos. Trans. A Math. Phys. Eng. Sci., 365(1857), 2053-2075, 2007.

van Andel, T.H. and Davies, W. Neanderthals and Modern Humans in the European Landscape of the Last Glaciation: Archaeological results of the stage 3 project. McDonald Institute for Archaeological Research Monographs, Cambridge, 2003.

Vita-Finzi, C. and Higgs, E. Prehistoric Economy in the mount Carmel area of Palestine: Site catchment analysis. P. Prehist. Soc., 36, 1970.

Vaittinada Ayar, P., Vrac, M., Bastin, S., Carreau, J., Déqué, M., and Gallardo, C. Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: Present climate evaluations. Clim Dyn, 46(3) 1301-1329, 2016.

Vrac, M., Marbaix, P., Paillard, D., and Naveau, P. Non-linear statistical downscaling of present and LGM precipitation and temperatures over Europe. Clim. Past, 3(4), 669-682, 2007.

Walter, H. and Box, E. Global classification of natural terrestrial ecosystems. Vegetatio, 32(2), 75-81, 1976.

Wolff, E. W., Chappellaz, J., Blunier, T., Rasmussen, S. O., and Svensson, A. Millennial-scale variability during the last glacial: The ice core record.Quaternary Sci. Rev., 29(21), 2828-2838, 2010.

Wood, A. W., Leung, L. R., SRidhar, V., and Lettenmaier, D. P. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, Clim. Change, 62, 189–216, 2004.

Wood, S. N. Modelling and Smoothing Parameter Estimation with Multiple Quadratic Penalties, J. R. Stat. Soc. B, 62(2), 413-428, 2000.

Wood, S. N. Stable and efficient multiple smoothing parameter estimation for generalized additive models, J. Am. Stat. Assoc., 99(467), 2004.

Wood, S. N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J. R. Stat. Soc. (B) 73(1), 3-36, 2011.

Wu, H., Guiot, J., Brewer, S., and Guo, Z. Climatic changes in Eurasia and Africa at the last glacial maximum and mid-Holocene: reconstruction from pollen data using inverse vegetation modelling, Clim. Dynam., 29(2-3), 211-229, 2007.

Yokoyama, Y., Lambeck, K., De Deckker, P., Johnston, P., and Fifield, L. K. Timing of the Last Glacial Maximum from observed sea-level minima. Nature, 406(6797), 713-716, 2000.

**Supprimé:** :

**Supprimé:** Vicente-Serrano, S.M., Beguería, S., and López-Moreno, J.I. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. J. Climate, 23(7):1696-1718, 2010.

**Supprimé:** S. Generalized additive models: an introduction with R. Chapman and Hall/CRC, Boca Raton, FL, 2006

**Table 1. Model selection among all possible combinations of variables for the temperature and the precipitations. Only the five models with the lowest AIC are presented. AIC scores, differences in AIC compared to the lowest scoring model ($\Delta_{AIC}$), and AIC weights ($w_{AIC}$ = exp(-0.5 × $\Delta_{AIC}$) ∈ [0,1], representing the relative likelihood of the models) are reported.**

| Candidate model | AIC | $\Delta_{AIC}$ | $w_{AIC}$ |
|---|---|---|---|
| Bilinear interpolation | | | |
| *Temperature* | | | |
| T+elv+Aco | 869468.4 | 0 | 1 |
| T+elv | 878862.9 | 9394.52 | 0 |
| T+Aco | 951357.5 | 81889.14 | 0 |
| T | 955986.9 | 86518.50 | 0 |
| elev+Aco | 1605682.9 | 736214.55 | 0 |
| *Precipitations* | | | |
| T+P+elev+Aco+RH+SLP | 275068.7 | 0 | 1 |
| T+P+elev+Aco+SLP | 281222.3 | 6153.56 | 0 |
| T+P+elev+Aco+RH | 283266.4 | 8197.70 | 0 |
| T+P+elev+RH+SLP | 288574.6 | 13505.85 | 0 |
| T+P+elev+Aco | 288864.1 | 13795.35 | 0 |
| Bicubic interpolation | | | |
| *Temperature* | | | |
| T+elv+Aco | 896517.0 | 0 | 1 |
| T+elv | 904724.6 | 8207.52 | 0 |
| T+Aco | 952109.7 | 55592.63 | 0 |
| T | 955079.5 | 58562.43 | 0 |
| elev+Aco | 1605402.4 | 708885.36 | 0 |
| *Precipitations* | | | |
| T+P+elev+Aco+RH+SLP | 274137.9 | 0 | 1 |
| T+P+elev+Aco+SLP | 278178.8 | 4040.881 | 0 |
| T+P+elev+Aco+RH | 283162.3 | 9024.412 | 0 |
| T+P+elev+Aco | 286857.5 | 12719.589 | 0 |
| T+P+elev+RH+SLP | 286926.3 | 12788.379 | 0 |
| Kriging | | | |
| *Temperature* | | | |
| T+elv+Aco | 878970.1 | 0 | 1 |
| T+elv | 888326.4 | 9356.24 | 0 |
| T+Aco | 952033.1 | 73062.92 | 0 |
| T | 956903.8 | 77933.70 | 0 |
| elev+Aco | 1607626.2 | 728656.08 | 0 |
| *Precipitations* | | | |
| T+P+elev+Aco+RH+SLP | 269767.4 | 0 | 1 |
| T+P+elev+Aco+SLP | 276171.0 | 6403.59 | 0 |
| T+P+elev+Aco+RH | 277692.0 | 7924.60 | 0 |
| T+P+elev+Aco | 283099.2 | 13331.79 | 0 |
| T+P+Aco+RH+SLP | 286495.6 | 16728.23 | 0 |

Tableau mis en forme

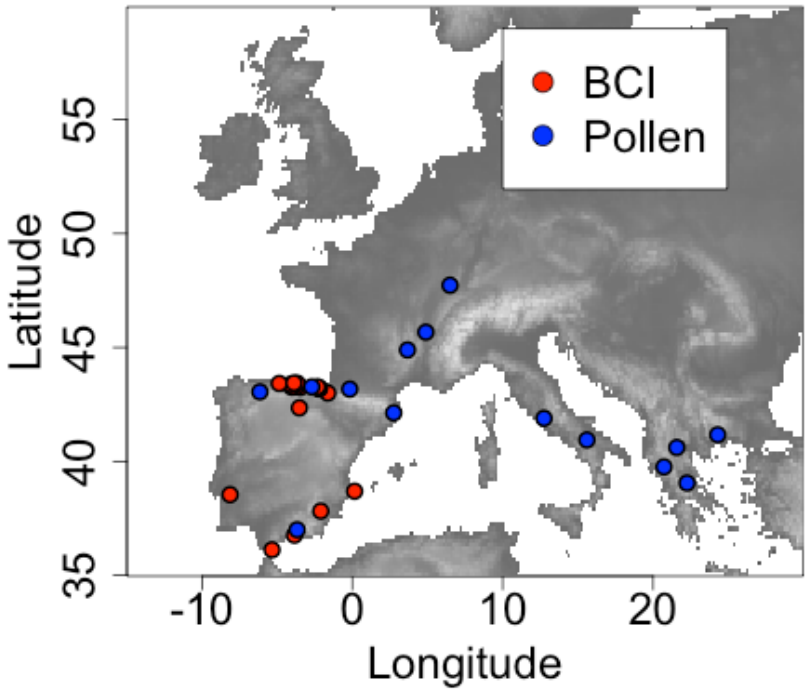
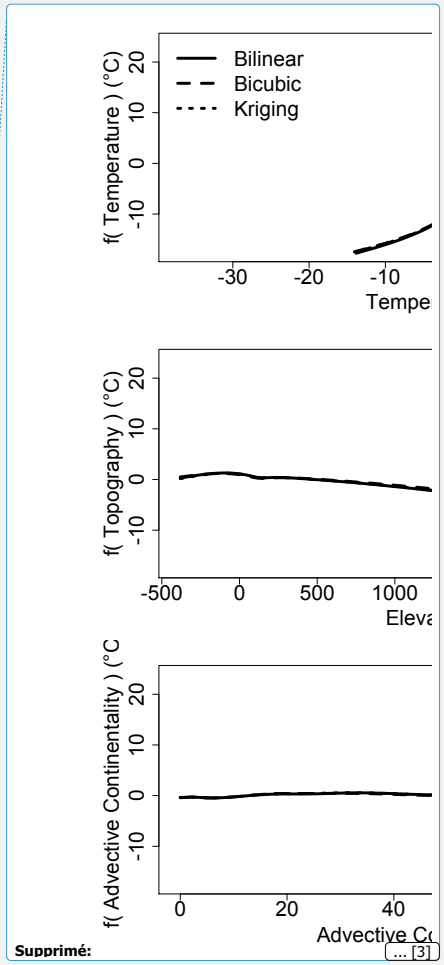



Figure 1. Study area and locations of the sites used for reconstructing local climate variables estimated on the basis of pollen and vertebrate fossils, used for the evaluation of the method. The grey scale represents elevation.




Supprimé: ... [3]

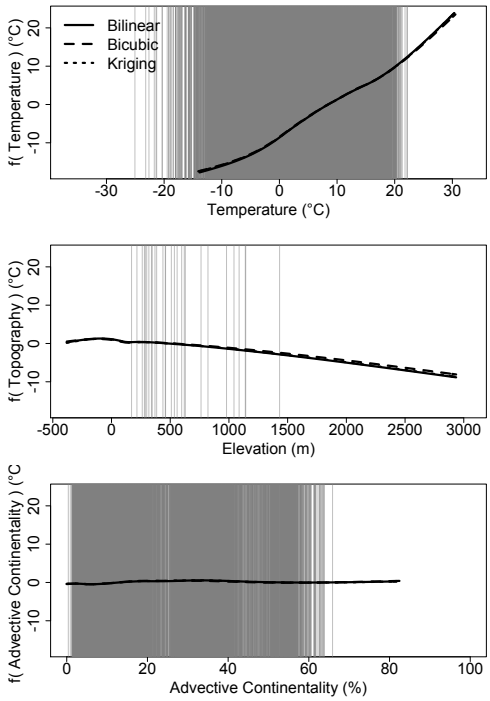
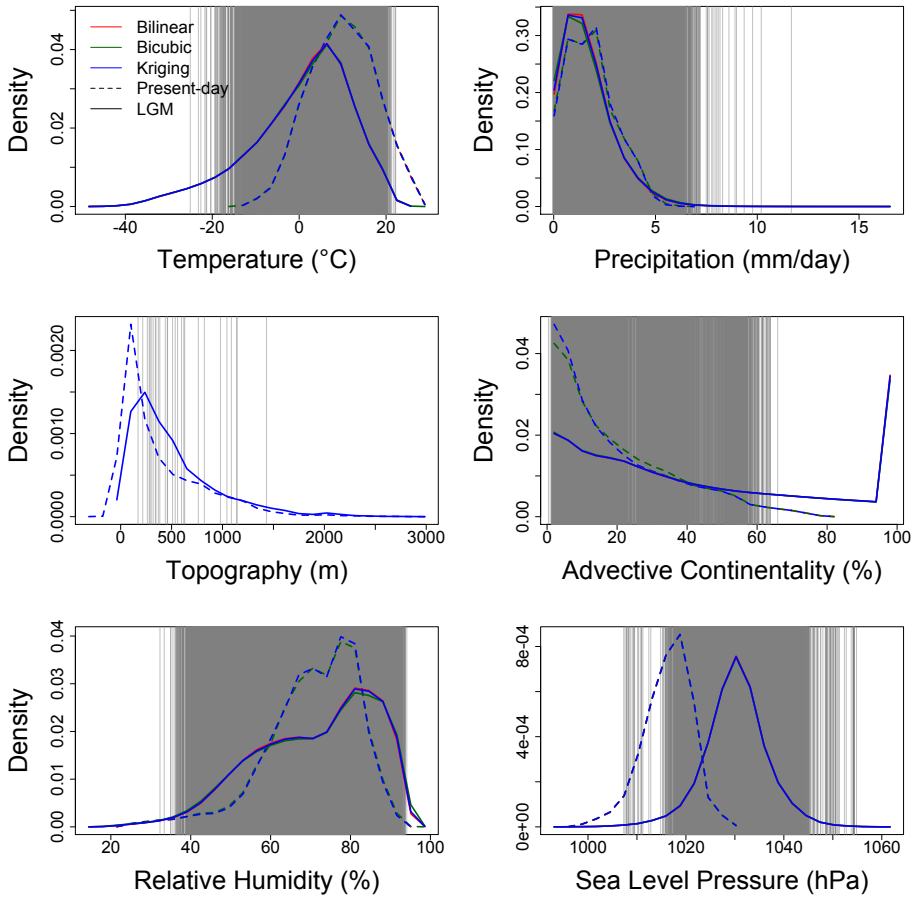



**Figure 2. Splines of the GAM for temperature. The splines are scaled to the same range to allow for visual estimation of their relative importance. The range of the x-axes combines the ranges of values for the present-day period and the LGM. The grey lines indicate the values for the 12 months over the 50 years during the LGM at the archaeological sites of Figure 1 (except for the elevation, for which there is only one value per site).**



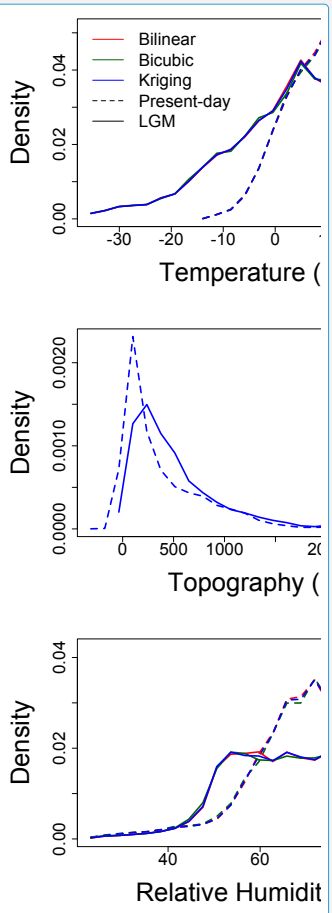




**Figure 3. Histograms of the predictor variables for the present-time (1961-1990; dashed lines) and for the LGM (solid lines) over Western Europe, using the bilinear (red), bicubic (green) and kriging (blue) interpolations. The grey lines indicate the values for the 12 months over the 50 years during the LGM at the archaeological sites of Figure 1 (except for the elevation, for which there is only one value per site).**

Supprimé:

Supprimé: 2

Supprimé: 1960

Supprimé: solid

Supprimé: dashed

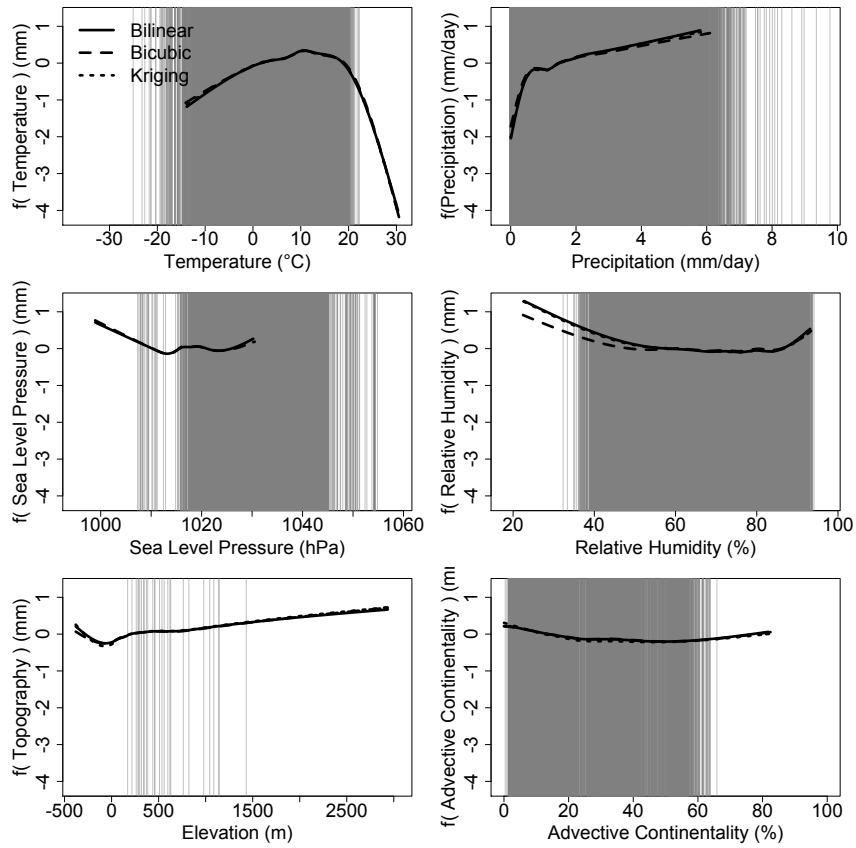
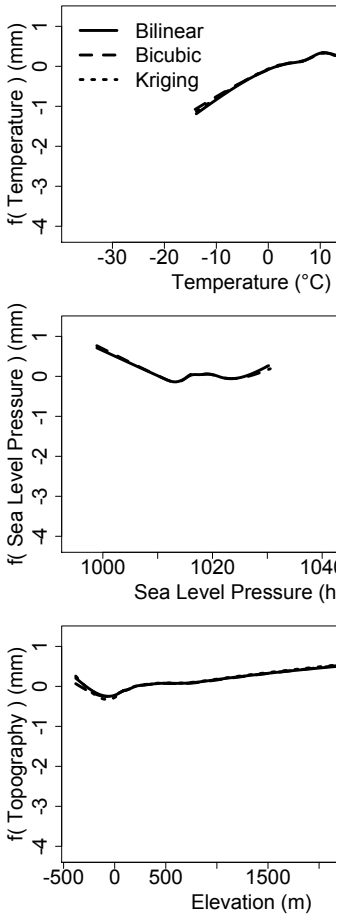
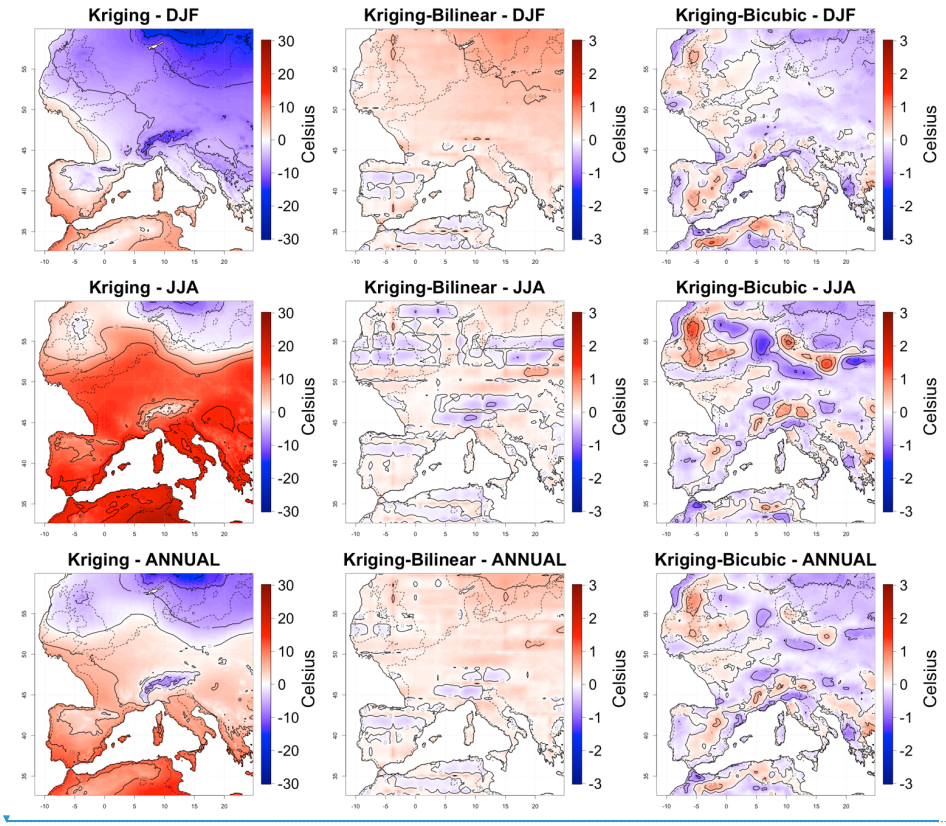


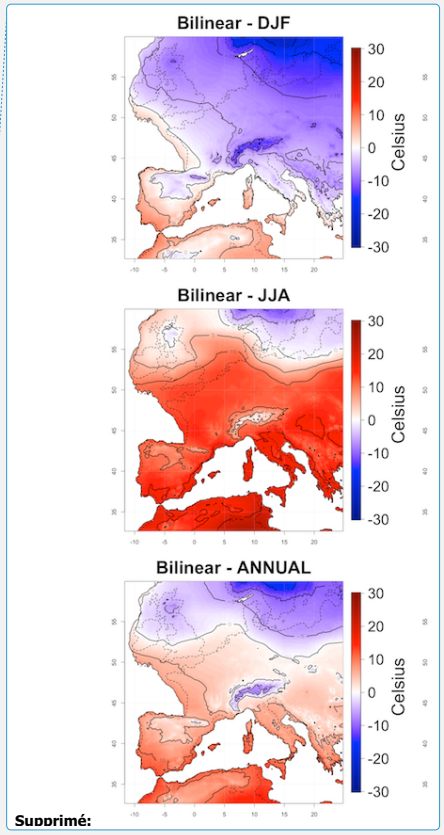

**Figure 4. Splines of the GAM for precipitations. The splines are scaled to the same range to allow for visual estimation of their relative importance.** The range of the x-axes combines the ranges of values for the present-day period and the LGM. The grey lines indicate the values for the 12 months over the 50 years during the LGM at the archaeological sites of Figure 1 (except for the elevation, for which there is only one value per site).




Supprimé:

Supprimé: 3




**Supprimé:**

**Figure 5. Mean distributions of monthly mean downscaled temperatures over Western Europe during the LGM for winter (December, January, February), summer (June, July, August), and the whole year, computed over 50 years for the kriging interpolation technique, and difference between the kriging and the other two techniques. Downscaling was performed for each month independently, but results are combined into seasons to summarise the results.**

**Supprimé:** 4. Maps of mean

**Supprimé:** three interpolation techniques

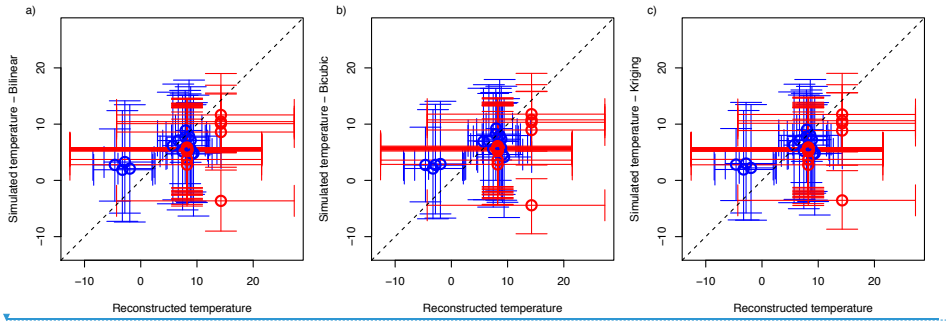
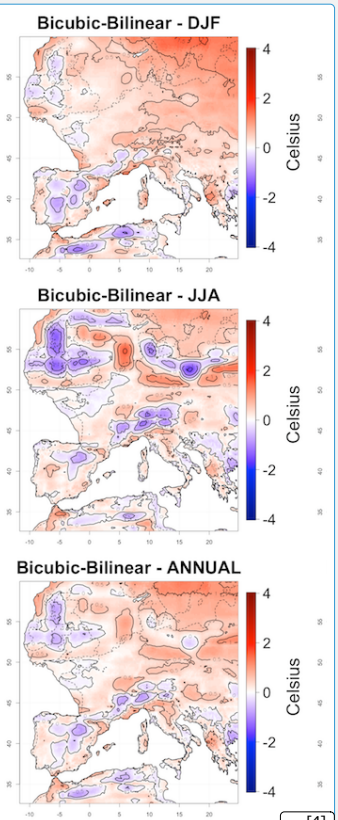



**Figure 6.** Comparison of reconstructed vs. downscaled temperatures for the LGM based on the BCI indices (red), and from Wu et al. (2007)'s reconstructions (blue) for a) the bilinear, b) the bicubic and c) the kriging interpolations. The circles represent the mean temperature values for the two reconstruction methods (x-axis) and the downscaled values over the 50 simulated years (y-axis). The horizontal error bars represent the range of temperature values for the reconstruction method (minimum and maximum over the whole zonobiome for the BCI indices, mean temperature of the coldest and warmest month for Wu et al. 2007). The vertical error bars correspond to the mean temperature of the coldest and warmest month over the 50 simulated years.



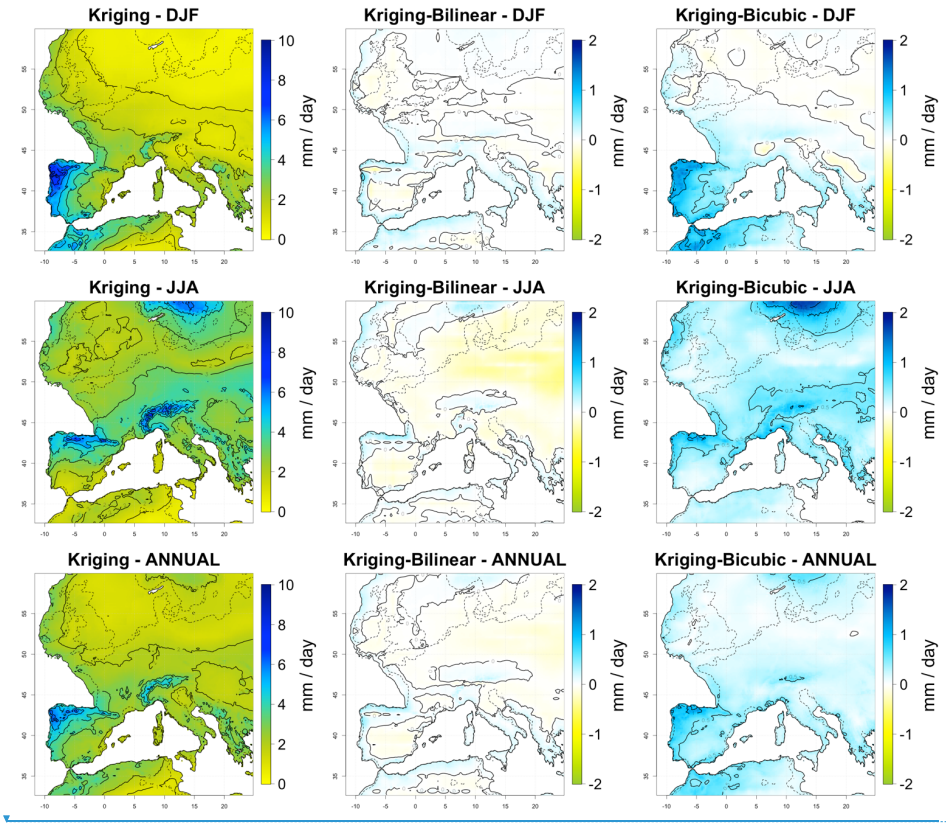

Supprimé: ... [4]

Supprimé: Boxplot

Supprimé: , simulated

Supprimé: a, c, e

Supprimé: , d, f).



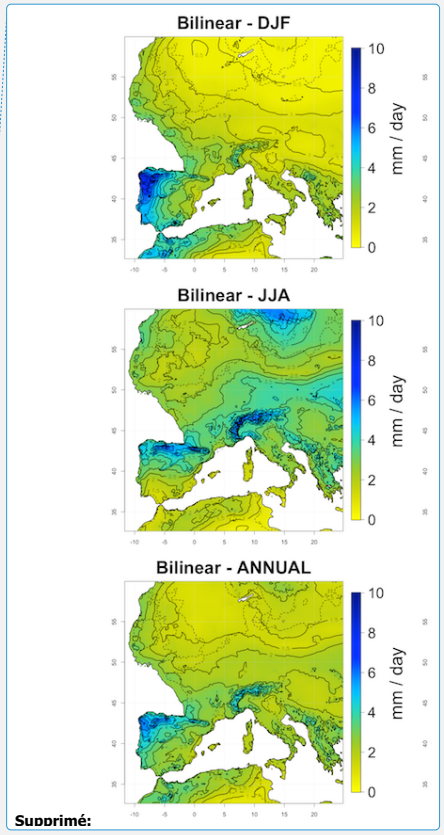




**Figure 7.** Mean distributions of downscaled daily precipitations over Western Europe during the LGM for winter (December, January, February), summer (June, July, August), and the whole year, computed over 50 years for the kriging interpolation technique, and difference between the kriging and the other two techniques. Downscaling was performed for each month independently, but results are combined into seasons to summarise the results.

Supprimé:

Supprimé: Maps of mean

Supprimé: three interpolation techniques

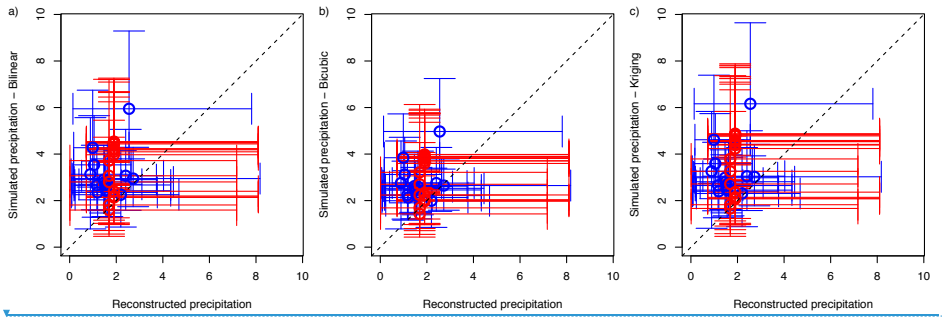
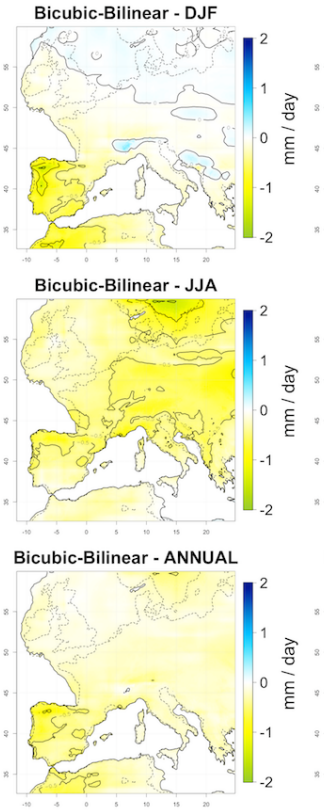



**Figure 8. Boxplot of reconstructed vs. downscaled precipitations for the LGM based on the BCI indices (red), and from Wu et al. (2007)'s reconstructions (blue) for a) the bilinear, b) the bicubic and c) the Kriging interpolations. The circles represent the mean temperature values for the two reconstruction methods (x-axis) and the downscaled values over the 50 simulated years (y-axis). The horizontal error bars represent the range of precipitation values for the reconstruction method (minimum and maximum over the whole zonobiome for the BCI indices, mean precipitation of the coldest and warmest month for Wu et al. 2007). The vertical error bars correspond to the mean precipitation of the driest and wettest month over the 50 simulated years.**



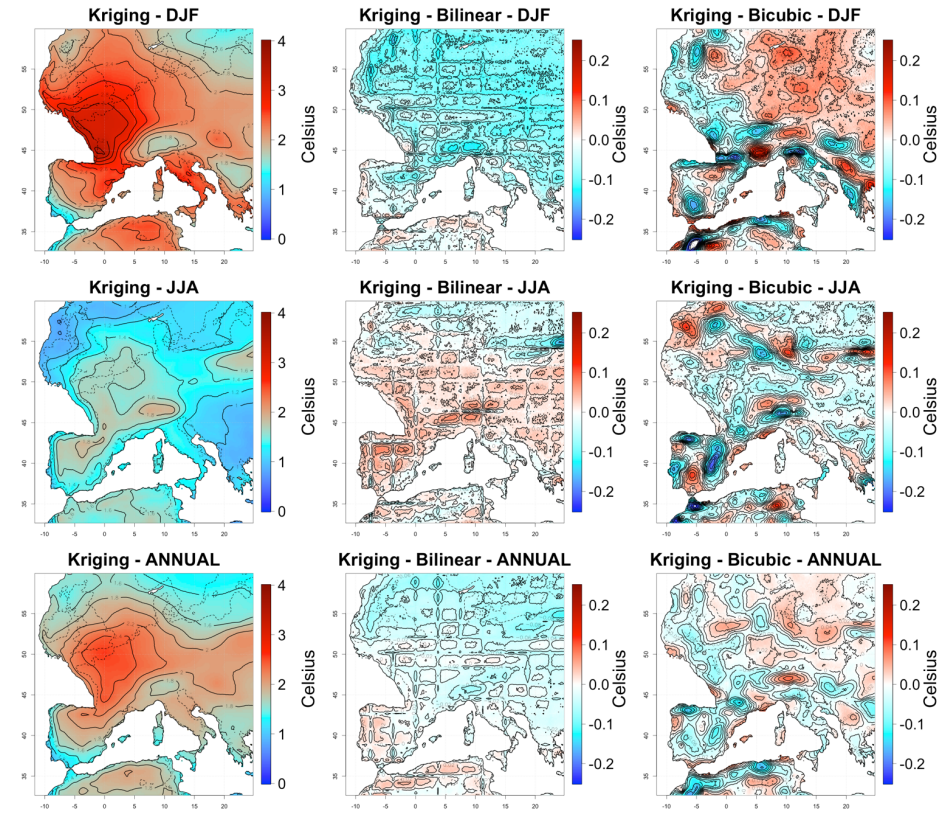

**Supprimé:**

**Supprimé:** Maps of difference in mean distributions of downscaled daily precipitations over Western Europe during the LGM for winter (December, January, February), summer (June, July, August), and the whole year, computed over 50 years, between the three interpolation techniques. ... [5]

**Supprimé:** , simulated

**Supprimé:** a, c, e

**Supprimé:** , d, f).



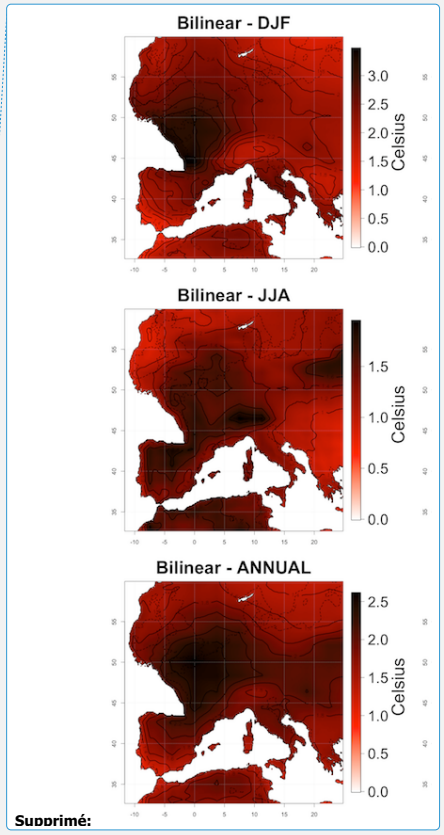




**Figure 9. Maps of temporal variations (standard deviation of each month across 50 years) of downscaled monthly mean temperatures over Western Europe during the LGM averaged over winter (December, January, February), summer (June, July, August), and the whole year for the kriging interpolation technique, and difference between the kriging and the other two techniques. Variability was computed for each month independently, but results are combined into seasons to summarise the results.**

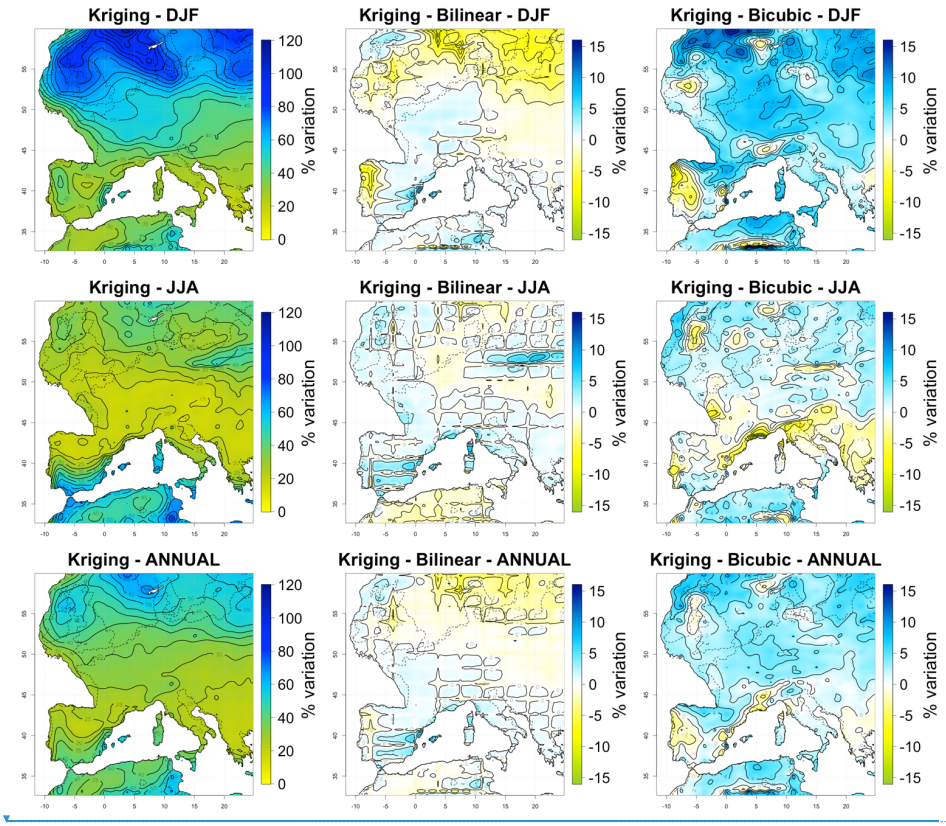
Supprimé:

Supprimé: 10

Supprimé: three

Supprimé: Color scales differ between maps

Supprimé: better distinguishing the spatial artifacts



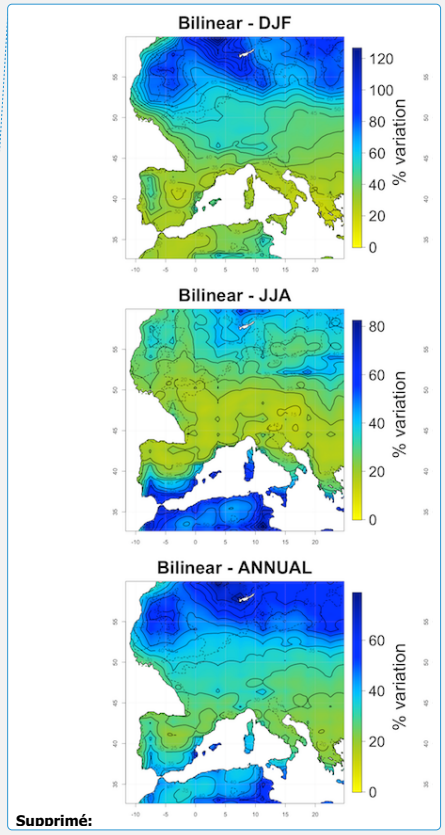




**Figure 10. Maps of temporal variations (coefficient of variation of each month across 50 years) of downscaled daily precipitations over Western Europe during the LGM averaged over winter (December, January, February), summer (June, July, August), and the whole year, computed over 50 years for the kriging interpolation technique, and difference between the kriging and the other two techniques. Variability was computed for each month independently, but results are combined into seasons to summarise the results.**

Supprimé: 11

Supprimé: three

Supprimé: Color scales differ between maps

Supprimé: better distinguishing the spatial artifacts

**Page 12 : [1] Supprimé** **Guillaume Latombe** **01/04/2018 17:22:00**

The three interpolation techniques generated similar STI and SPI surfaces (Fig. S19). The main difference occurred for the SPI from the bicubic interpolation, which showed less months with "normal" precipitation (i.e. either drier or wetter than baseline) than the other two interpolation techniques in the center and North of Europe. As for the temporal variability, this result is due to the fact that, contrary to the other two interpolation techniques, bicubic interpolation generates values outside of the original coarse-scale values, and is therefore more likely to over- or underestimate them.

**Page 18 : [2] Déplacé vers la page 18 (Déplacement n°3)** **Guillaume Latombe** **01/04/2018 17:22:00**

Jennings, R., Finlayson, C., Fa, D., and Finlayson, G. Southern Iberia as a refuge for the last Neanderthal populations. J. Biogeogr., 38(10), 1873-1885, 2011.

Jiménez-Espejo, F. J., Martínez-Ruiz, F., Finlayson, C., Paytan, A., Sakamoto, T., Ortega-Huertas, M., Finlayson, C., Ijima, K., Gallego-Torres, D., and Fa, D. Climate forcing and Neanderthal extinction in Southern Iberia: insights from a multiproxy marine record. Quaternary Sci. Rev., 26(7), 836-852, 2007.

**Page 23 : [3] Supprimé** **Guillaume Latombe** **01/04/2018 17:22:00**

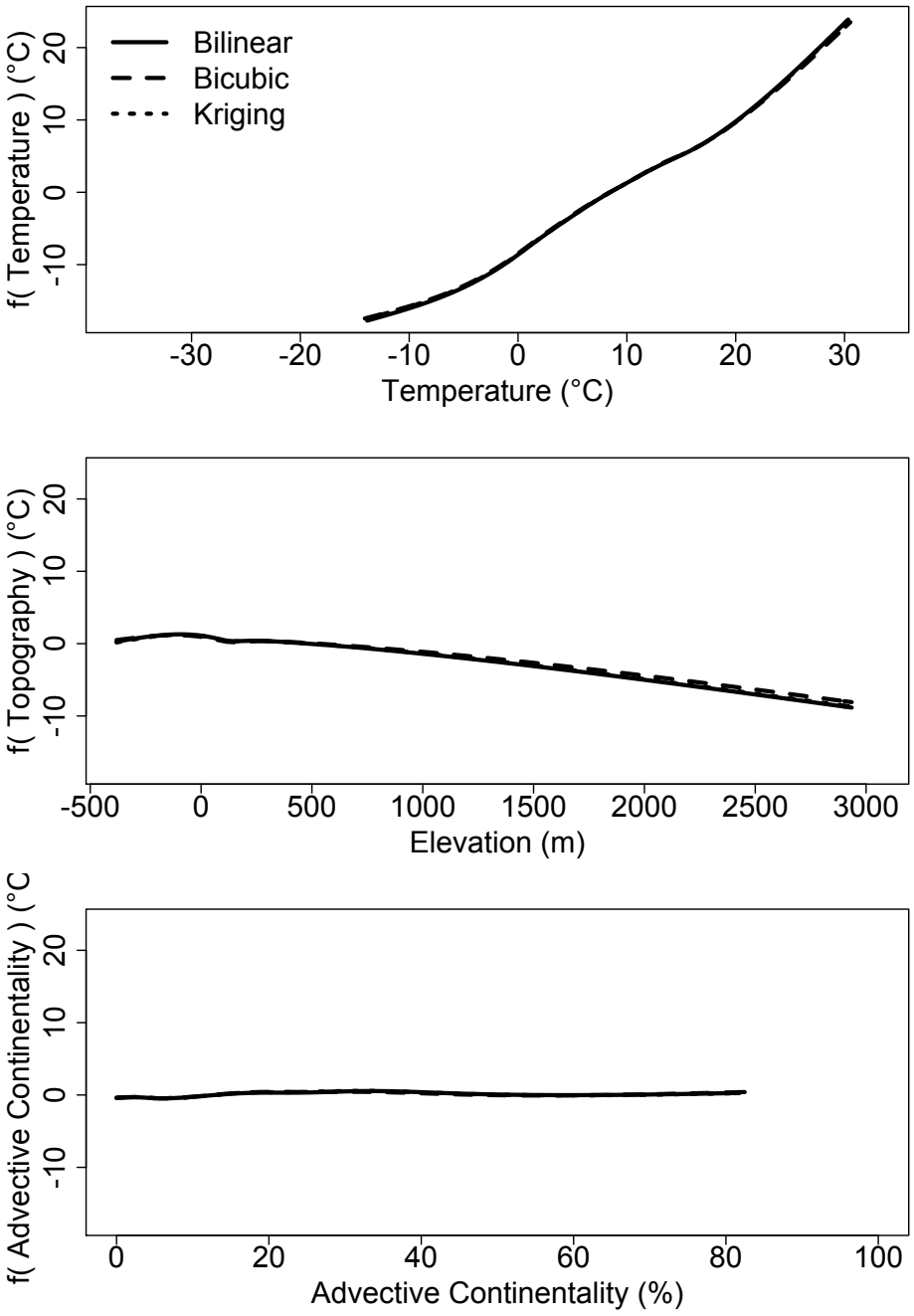



**Figure 1.**

| Page 28 : [4] Supprimé | Guillaume Latombe | 01/04/2018 17:22:00 |
|---|---|---|

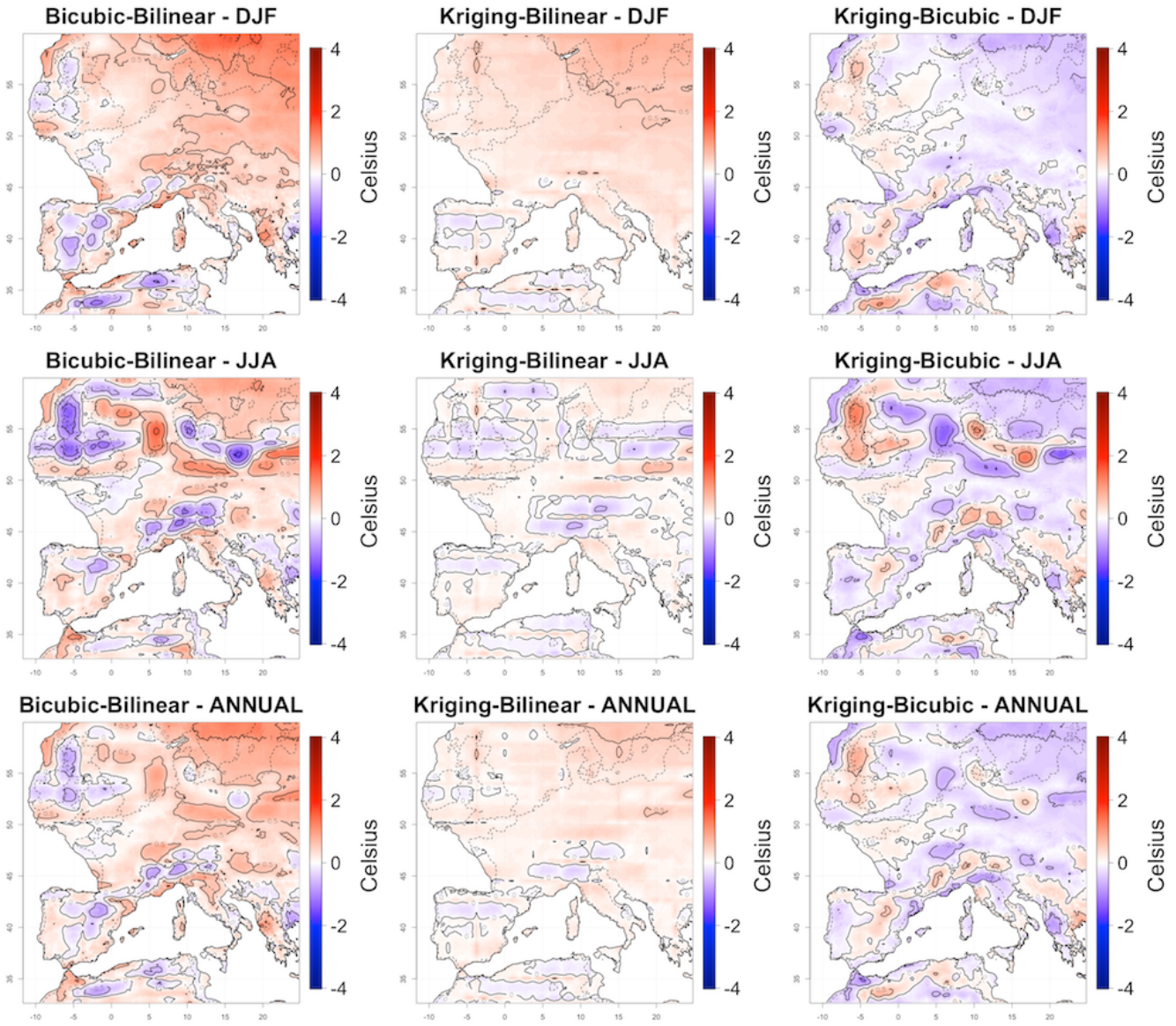
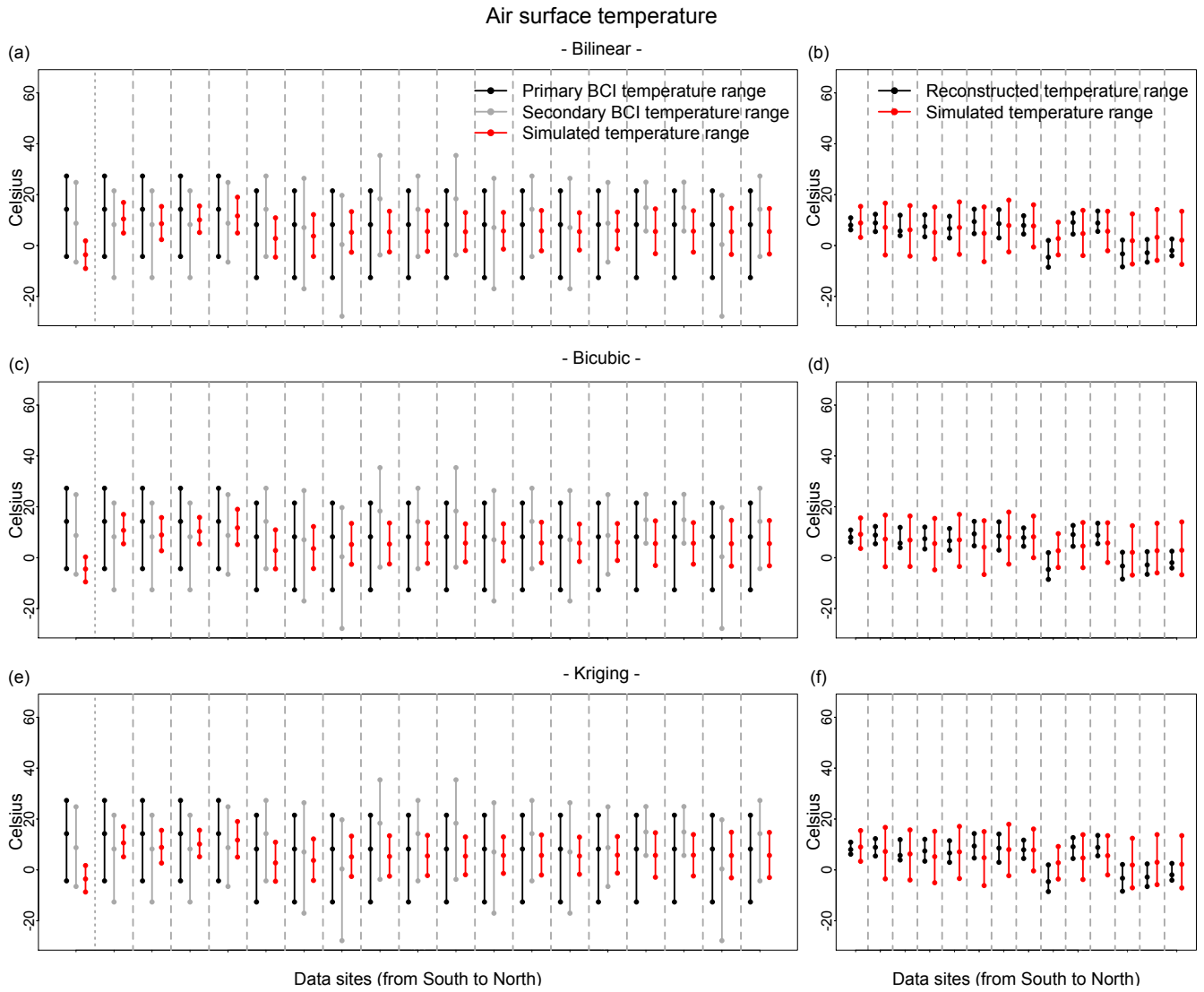



**Figure 5. Maps of difference in mean distributions of monthly mean downscaled temperatures over Western Europe during the LGM for winter (December, January, February), summer (June, July, August), and the whole year, computed over 50 years, between the three interpolation techniques.**




**Page 30 : [5] Supprimé** **Guillaume Latombe** **01/04/2018 17:22:00**

**Maps of difference in mean distributions of downscaled daily precipitations over Western Europe during the LGM for winter (December, January, February), summer (June, July, August), and the whole year, computed over 50 years, between the three interpolation techniques.**

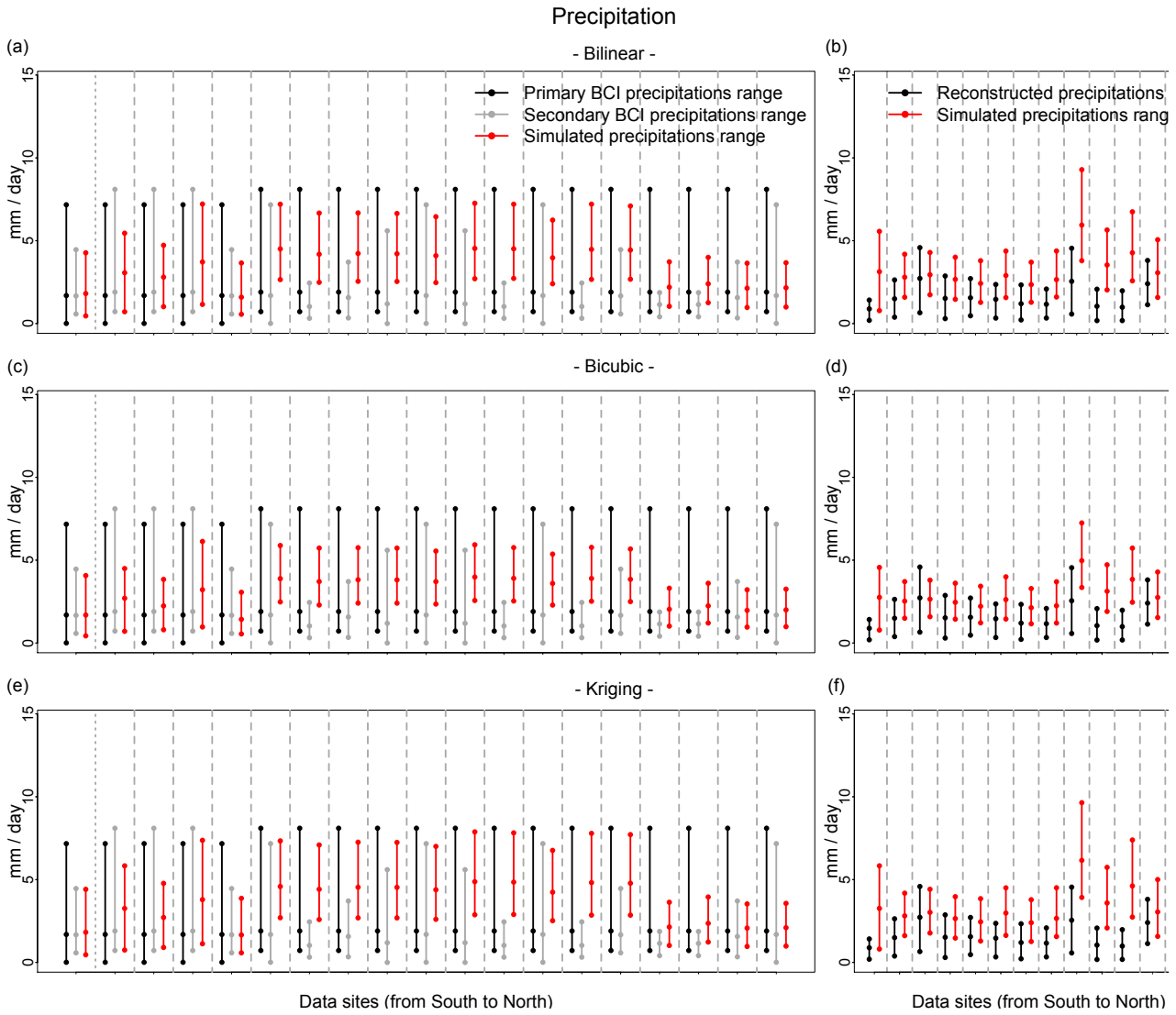

Precipitation
(a)
- Bilinear -
(b)
Primary BCI precipitations range
Secondary BCI precipitations range
Simulated precipitations range
Reconstructed precipitations
Simulated precipitations rang
(c)
- Bicubic -
(d)
(e)
- Kriging -
(f)
mm / day
0
5
10
15
Data sites (from South to North)
Data sites (from South to North)

**Figure 9.**