Collated responses to reviews

Reviewer 1

Comment: clarification of data availability

Response: As the reviewer points out, the pollen-based reconstructions and the climate model simulations underpinning our reconstruction are in the public domain, and the data assimilation methodology is described in detail in another publication. The general approach used for the CO2 corrections, which the reviewer describes as a significant contribution, was published in Prentice et al. (2017) – although we provide the equations for the implementation of this approach in the current paper in Appendix 1. Therefore, the new results in this paper are indeed the global maps of reconstructed climate variables. These data are archived and will be made publicly available – however, we realise that it may not have been obvious that the citation to Cleator et al. (2019b) represented the reconstruction data set. We have modified the last sentence of the abstract to make it clear that the reconstruction data are available as follows:

Thus, the new reconstructions provides a benchmark created using clear and defined mathematical procedures that can be used for evaluation of the PMIP4/CMIP6 entry-card LGM simulations and are available at DOI:10.17864/1947.229

We have included a Data Availability section in the revised paper:

Data availability. The gridded data for the LGM reconstructions are available from http://dx.doi.org/10.17864/1947.229; the code used to generate these reconstructions is available from (10.5281/zenodo.3445166).

Comment: Varying definitions of the LGM

Response: The reviewer indicates that the definition of the LGM used in our paper (21±1 ka) differs from the interval used by Annan and Hargreaves of 21±2ka, and there is recent work on sea level (Ishiwa et al. 2019) which suggests the 'real' LGM was 19.1-19.7 ka, with a plateau prior to this from 20.4-25.9ka. Our choice of this time interval reflects the fact that the LGM is conventionally defined in PMIP at 21 ka and most of the pollen-based reconstructions of this interval included in the Bartlein et al data set are from the 21±1 ka. We are aware that there is still controversy over the timing of the LGM, with both younger and older ages mooted for the actual maximum ice volume/sea-level lowering (see e.g. Peltier and Fairbanks, 2006; Clark et al., 2009; Lambeck et al., 2014). Even the recent work by Ishiwa et al. (2019) points out that the sea level drop after 19.7 ka was only 10m and that there was a long plateau with stable low sea level prior to this and encompassing the 21 ka interval. Since our aim is to produce a data set for benchmarking new PMIP LGM simulations, which will be run with boundary conditions for 21 ka (Kageyama et al., 2017), the exact date of the LGM is therefore not an issue. However, we agree that there is a difference between the true definition of the LGM and the convention used for modelling purposes, and that this is not clear from our introductory text, so we have expanded our definition (line 55 onwards) as follows:

At the Last Glacial Maximum (LGM, conventionally defined for modelling purposes as 21 000 years ago), insolation was quite similar to the present, but global ice volume was at a maximum, eustatic sea level was close to a minimum, long-lived greenhouse gas concentrations were lower and atmospheric aerosol loadings higher than today, and land surface characteristics (including vegetation distribution) were also substantially different from today.

Comment: Consider how to use new studies published after Bartlein et al.

Response: The reviewer points out that we refer to several new studies since the Bartlein paper on which the analysis is based, and there are more, and that it would be nice to think these could be assimilated into a future dataset to maybe close some of the large 'no data' holes in the results. We thoroughly agree that it would be good to plug the gaps, and this will be an effort for the future. The three papers that we cite at lines 361-363 (Flantua et al., 2015; Herbert and Harrison, 2016; Harrison et al., 2016) demonstrate that there are pollen records available that would plug the gaps, but alas do not provide quantitative reconstructions at these sites. The Izumi and Bartlein (2016) paper provides an inversion-based reconstruction for North American – this region is already relatively well covered in the Bartlein et al data set. Similarly Mauri et al. (2015) provide a new gridded reconstruction for Europe – again a region that is well covered in the Bartlein et al data set. However, we are aware of new pollen-based quantitative reconstructions embracing the LGM for individual sites (e.g. in Africa, China, Russia, southern Europe) and compiling these reconstructions would certainly be a worthwhile effort in the future. Our method also lends itself to combining pollen-based reconstructions with other quantitative estimates of terrestrial palaeoclimate, and again this should be something that is done in the future. We have revised the paragraph describing future possibilities to expand the current data set to spell out some of these opportunities more clearly (lines 304-309), as follows:

Some areas are still poorly covered by quantitative pollen-based reconstructions of LGM climate, most notably South America. More pollen-based climate reconstructions would provide one solution to this problem – and there are many pollen records that could be used for this purpose (Flantua et al., 2015; Herbert and Harrison, 2016; Harrison et al., 2016). There are also quantitative reconstructions of climate available from individual sites (e.g. Lebamba et al., 2012; Wang et al., 2014; Loomis et al., 2017; Camuera et al., 2019) that should be incorporated into future data syntheses. It would also be possible to incorporate other sources of quantitative information, such as chironomid-based reconstructions (e.g. Chang et al., 2015) within the variational data assimilation framework.

Additional references

Camuera, J., Jiménez-Moreno, G., Ramos-Román, M.J., García-Alix, A., Toney, J.L., Anderson, R.S., Jiménez-Espejo, F., Bright, J., Webster, C., Yanes, Y., José S. Carrión, J.S., 2019. Vegetation and climate changes during the last two glacial-interglacial cycles in the western Mediterranean: A new long pollen record from Padul (southern Iberian Peninsula), Quaternary Science Reviews, 205, 86-105, https://doi.org/10.1016/j.quascirev.2018.12.013.

Chang, J.C., Shulmeister, J., Woodward, C., Steinberger, L., Tibby, J., Cameron Barr, C., 2015. A chironomid-inferred summer temperature reconstruction from subtropical Australia during the last glacial maximum (LGM) and the last deglaciation, Quaternary Science Reviews, 122, 282-292, https://doi.org/10.1016/j.quascirev.2015.06.006.

Lebamba, J., Vincens, A., and Maley, J.: Pollen, vegetation change and climate at Lake Barombi Mbo (Cameroon) during the last ca. 33 000 cal yr BP: a numerical approach, Clim. Past, 8, 59-78, https://doi.org/10.5194/cp-8-59-2012, 2012.
Loomis, S. E., Russell, J. M., Verschuren, D., Morrill, C., De Cort, G., Sinninghe Damsté, J. S., … Kelly, M. A. (2017). The tropical lapse rate steepened during the Last Glacial Maximum. *Science advances*, *3*(1), e1600815. doi:10.1126/sciadv.1600815
Wang, Y., Herzschuh, U., Shumilovskikh, L. S., Mischke, S., Birks, H. J. B., Wischnewski, J., Böhner, J., Schlütz, F., Lehmkuhl, F., Diekmann, B., Wünnemann, B., and Zhang, C.: Quantitative reconstruction of precipitation changes on the NE Tibetan Plateau since the Last Glacial Maximum – extending the concept of pollen source area to pollen-based climate reconstructions from large lakes, Clim. Past, 10, 21-39, https://doi.org/10.5194/cp-10-21-2014, 2014.

Comment: L59: change to 'lower, atmospheric aerosol. . .'
Response: We have made this change.

Comment: L321: comma after 'however'
Response: We have made this change (now line 274).

Response to Review 2

Comment: Mathematical details of the technique applied in this study are available in Cleator et al., 2019a. This is an arXiv preprint. It is not clear whether the latter is intended for a peer-reviewed journal or whether it was part of a thesis examination.

Response: This article is now accepted for publication in the Journal of Advances in Modeling Earth Systems. As a result of the review process for JAMES, it was pointed out that the Gaussian correlation function we used did not yield a full rank matrix; we have therefore moved to using a modified Bessel function that closely matches the behaviour of the original Gaussian function and yields a correlation matrix that is full rank and positive. We have checked that this change does not make a substantial difference to the global reconstructions presented here (although it changes some numbers slightly, and we have amended the text to reflect this) and does not change the conclusions of our paper. The revised method paper is available from arXiv (arXiv:1902.04973v2, https://arxiv.org/abs/1902.04973v2) and will soon be available in JAMES. We have updated the figures and the text here to reflect the use of the revised function.

Comment: There are a number of points in this approach which deserve discussion. For this reason it would have been better to see these method details in the Climate of the Past paper, so that the paper, the review, and possible responses constitute a self-contained contribution.

Response: We wanted to focus the discussion here on the results (i.e. the reconstructions of LGM climate) rather on the mathematical details of the method. These details will shortly be available in JAMES and are given in the pre-print article. However, we agree that it would be worth expanding the section on the application of the data assimilation method and the choice of length scales (section 2.4) to provide more details. We have modified this section as follows:

Variational data assimilation techniques provide a way of combining observations and model outputs to produce climate reconstructions that are not exclusively constrained to one source of information or the other (Nichols, 2010). We use the 3D-variational method, described in Cleator et al. (2019a) to find the maximum a posteriori estimate (or analytical reconstruction) of the palaeoclimate given the site-based reconstructions and the model-based prior. The method constructs a cost function, which describes how well a particular climate matches both the site-based reconstructions and the prior, by assuming the reconstructions and prior have a Gaussian distribution. To avoid sharp changes in time and/or space in the analytical reconstructions, the method assumes that the prior temporal and spatial covariance correlations are derived from a modified Bessel function, in order to create a climate anomaly field that is smooth both from month to month and from grid cell to grid cell. The degree of correlation is controlled through two length scales: a spatial length scale that determines how correlated the covariance in the prior is between different geographical areas, and a temporal length scale that determines how correlated it is through the seasonal cycle. The site-based reconstructions are assumed to have negligible correlations at these space and time scales. The maximum a posteriori estimate is found by using the limited memory Broyden- Fletcher-Goldfarb-Shanno method (Liu and Nocedal 1989) to determine the climate that minimises the cost function. A first order estimate of the analysis uncertainty covariance is also computed.

An observation operator based on calculations of the direct impact of [CO2] on water-use efficiency (section 2.3) is used in making the analytical reconstructions. The prior is constructed as the average of eight LGM climate simulations (section 2.2). We use an ensemble of different model responses to the same forcing to provide a series of physically consistent possible states, which can be viewed as perturbed responses and provide the variance around the climatology provided by the ensemble average. The prior uncertainty correlations are based on a temporal length scale (Lt) of 1 month and a spatial length scale (Ls) of 400km. Cleator et al., (2019a) have shown that a temporal length scale of 1 month provides an adequately smooth solution for the seasonal cycle, both using single sites and over multiple grid cells, as shown by the sensitivity of the resolution matrix (Menke, 2012; Delahaies et al., 2017) to changes in the temporal length scale. Consideration of the spatial spread of variance in the analytical reconstruction shows that a spatial length scale of 400km also provides a reasonable reflection of the large-scale coherence of regional climate change.

Additional references:

Liu, D. C., & Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. Mathematical Programming, 45 (1), 503–528. doi: 10.1007/BF01589116

Delahaies, S., Roulstone, I., & Nichols, N. (2017). Constraining DALECv2 using multiple data streams and ecological constraints: analysis and application. Geoscientific Model Development (Online), 10 (7). doi: 10.5194/gmd-10-2635- 2017

Menke, W. (2012). Geophysical data analysis: Discrete inverse theory (Matlab 3rd ed.). Cambridge, Massachusetts: Academic Press.

Comment: Unlike what (roughly) obtains when using time series of a numerical weather prediction system, there is a priori no guarantee that the covariance matrix of a multi-model ensemble produces modes which satisfy “physical consistency”. Why would we expect that the inter-model differences provide knowledge about how different variables should co-vary?

Response: Our argument here is that the average response of all the models gives a measure of climatology. Numerical weather prediction uses ensembles of perturbed responses to provide a series of physically-consistent possible states, although there are examples of using multiple models to form an ensemble (see e.g. Johnson and Swinbank, 2009 - https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.383). Here we use an ensemble of different model responses to the same forcing, which can be viewed as producing perturbed responses to the general climatology. We have added a sentence in the method text (given above) to make this argument clearer.

Comment: In principle, a "prior" encodes what we a priori believe the climate could be. The authors have then chosen to mask regions with little update by observations, and leave visible the grid points where the observations have seriously shifted the prior. This seems at first sight reasonable because the idea is to focus on the pollen reconstructions and not on the PMIP3 output. Yet, at face value, this approach is inconsistent with a Bayesian interpretation. Grid points of strong update are associated, in the Bayesian interpretation, with a very small marginal likelihood (a wrong prior means a wrong model).

Response: We are starting from the assumption that the pollen-based reconstructions are more likely to be correct than the model simulations; but that the model simulations provide us with physically-consistent relationships in space and time (which cannot be obtained from the pollen). This comment is partly due to a misunderstanding about the masking, which is in fact determined by the variance rather than the absolute change. Only areas with an improved variance are shown (i.e. left unmasked). This means that the likelihood that these reconstructions represent the true climate is significantly improved from the prior. This only happens if the variance in the observations is small and the variance in the prior is big. By using both local and global measures of the variance in the prior, we avoid a situation where the variance in the prior is small but shows a different signal from the pollen-based reconstructions. We will expand the text to make this clearer (lines 265-271) as follows:

The reliability of the analytical reconstructions was assessed by comparing these composite covariances with the uncertainties in the prior. We masked out cells where the inclusion of site-based reconstructions does not produce an improvement of > 5% from the prior. Since this assessment is based on a change in the variance, rather than absolute values, this masking removes regions where there are no pollen-based reconstructions or the pollen-based reconstructions have very large uncertainties.

Comment: To what extent should we be concerned that the posterior variance remains influenced by the prior variance? Indeed, mathematically, the posterior variance is bounded by the prior variance, which if we admit the models are really off, is meaningless.

Response: The posterior variance is influenced (though not bounded or limited by) by the prior variance. However, since areas that have a small change to prior covariance are masked out, only areas with pollen-based reconstructions with low variance are used in the reconstruction. Hence, the prior variance only influences the posterior variance in areas that are well constrained by observations. Furthermore, since the prior variance is based partly on the global variance for each variable, the only way to have a large prior variance affecting the posterior variance is for all models to agree well globally and locally and the observation to have a low variance such that the posterior variance has improved upon the prior variance change by over the 5% cutoff. We agree that the choice of the cutoff is somewhat arbitrary (as we state in the discussion section, (lines 406-412), though guided by examination of the impact of this cutoff on the reconstructions, and that it would be useful to develop an objective way of determining an appropriate limit. We have expanded the discussion to suggest ways forward here (lines 406-412) as follows:

We have used a <5% reduction in the analytical uncertainty compared to prior uncertainty to identify regions where the incorporation of site-based data has a negligible effect on the prior as a way of masking out regions for which the observations have effectively no impact on the analytical reconstructions. The choice of a 5% cut-off is arbitrary, but little would be gained by imposing a more stringent cut-off at the LGM given that many regions are represented by few observations. A more stringent cut-off could be applied for other time intervals with more data. We avoid the use of a criterion based on the analytical reconstruction showing any improvement on the prior because this could be affected by numerical noise in the computation. Alternative criteria for the choice of cut-off could be based on whether the analytical reconstruction had a reduced uncertainty compared to the pollen-based reconstructions or could be derived by a consideration of the condition number used to select appropriate length scales.

Comment: To what extent the prior covariance (link between different variables) may still be trusted at all if the models are so wrong? This remark strengthens the original concern about the physical meaning of the covariance matrix, even when the prior is only mildly up-dated. What is the advantage of this approach over a mere Gaussian interpolation (flat climate prior), which in this case might turn out to be more reliable and free of the dubious claim of "physical consistency"?

Response: We acknowledge that the climate models may not be correct, for example because the LGM simulations do not include all of the necessary forcings or show weak responses to these forcings. However, analyses of the PMIP simulations indicate that while the models show differences of both magnitude and sign in some regions, the overall LGM to present change is broadly consistent with what we know from observations. It is worth pointing out that many of these regional problems are associated with model dynamics rather than thermodynamics, which suggests that the models can be used to ensure physical consistency between surface variables. We try to overcome the problem of "all models being consistent but wrong" at a regional scale by combining global and local uncertainties to produce the uncertainty on the prior. In revising the section describing the variational approach (see above), we have tried to make our logic clearer here.

Comment: Were the length scales tested by some form of cross-validation (e.g. leave-one-out), or were they merely chosen because they are a priori reasonable?

Response: We did not use cross-validation to evaluate the choice of length scales, but instead we based the choice of length scales on sensitivity experiments (as described in the arXiv pre-print). Effectively we ran a series of tests to see how different choices affect the resolution matrices and the condition number. We selected a spatial length scale that provided a reasonable reflection of the large-scale coherence of regional climate change and also ensured that the covariance matrix was well-conditioned for inversion, and a temporal length scale that limited overlap between successive months. The selected length scales seem reasonable; for example, the spatial scale corresponds to a situation where there is little overlap between data points assuming an average catchment size for the pollen records on which the original reconstructions were based. Similarly, the selected temporal length scale produces plausible-looking seasonal cycles of temperature. We have expanded the text describing the application of the variational method (see above) to clarify how the length scales were chosen based on these sensitivity tests and a post-hoc evaluation of plausibility.

Comment: The arXiv paper provides the definition of the moisture index. It should be repeated here (moisture index is currently introduced l. 297 without definition)

Response: The MI was defined at line 155. In the present context the reference to MI is inappropriate because the text refers to a generic control by moisture availability rather than a specific index. We also note there was a crucial comma missing in this sentence! We have altered the text here to read:

which is generally taken into account by process-based ecosystem models, but not by statistical models, using projected changes in vapour pressure deficit or some measure of plant-available water

Comment: The authors should consider providing a link to supporting code. The maps are currently provided as University of Reading dataset (with a doi) but its lifecycle is detached from the present contribution. A dataset consistent with the current Climate of the Past submission, reflecting a possible response to concerns of the reviewers, might best be included as supplementary information. Have we lodged code somewhere?

Response: The data used to generate the maps are lodged at the University of Reading repository, with a DOI. This allows external users to generate their own maps and their own analyses using the reconstructions. A revised version of these data, reflecting minor changes in the data as a consequence of using a Bessel function, has now been lodged at the repository. The two data sets are linked, so that external users are directed to the updated version of the data set. We do not envisage any changes to the data set as a result of review of this CoP submission, but if there are further changes to the data set then the current data set can be updated and again linked. Thus, the data provided in the repository are indeed constantly linked to the lifecycle of the product. The code used to generate the reconstructions has been lodged at Zenodo, and we have provided the DOI for this code in the revised ms. We have added a data availability section to the ms as follows:

Data availability: The gridded data for the LGM reconstructions are available from http://dx.doi.org/10.17864/1947.229; the code used to generate these reconstructions is available from (10.5281/zenodo.3445166).

Comment: It is important to distinguish the notion of variance from the notion of uncertainty. They are not synonymous. Variance describes the second momentum of a distribution; un-certainty is a reference to an identified lack of knowledge. Only when the distribution is assumed reflects our knowledge of a given quantity is it legitimate to identify both.

Response: We agree that the use of terminology here is inaccurate and we need to be more precise. However, uncertainty is not simply an identified lack of knowledge! It is also used to refer to the limits on the precision of knowledge (as in the case where we talk about the uncertainties attached to a pollen-based climate reconstruction, which are partly a function of our ability to define precise relationships with existing training data sets). We have corrected the ms throughout to ensure that we use variance and uncertainty appropriately. We have made the following specific changes: l.34 error changed to uncertainty l.134 error changed to uncertainty l.268 error changed to variances l.269 error changes to covariances l.272-278 error changed to variance l.282 error changed to variances l.283 error changed to variances (We have also changed this for Figure 3 in the caption list section) l.147 uncertainty change to variances

Figure 2 caption, uncertainty changed to variances Figure 3 caption, uncertainty changed to variances l.406 uncertainty changed to variance

Comment: Multi-model ensembles, in general, cannot be said to capture our knowledge of the state of climate at a given time. For this reason, I would argue not to call the PMIP3 covariance a “background uncertainty”.

Response: We agree that models are not the only source of information about the state of the climate at a given time, and indeed our approach makes the assumption that the pollen-based reconstructions are more likely to represent the true state of the climate. We agree that the models may be wrong because they do not include all the appropriate forcings, because the response to these forcings is too weak, or because of inappropriate treatment of key feedbacks. We also agree that not all models are equally good (or bad) and that in an ideal world a prior should be reconstructed based only on an ensemble of well-validated models. However, the point of using climate models here is to provide a way of deriving physically consistent relationships between climate variables, given that we do not have reconstructions of all of the seasonal variables everywhere. Furthermore, there are comparatively few LGM simulations available and using a more limited number of “more likely to be correct” simulations to create the prior (and estimate its variance) does not seem to be a good option. In the future, it might be possible to combine PMIP3 and PMIP4 simulations to create a more robust/plausible prior, but this is currently not possible.

Comment: The legend of Figure 2 clearly identifies “uncertainties” with “standard deviation of the non-dimensionalised multi-model ensemble” but this seems inadequate to me. Adding to the confusion, different qualifiers occur throughout the text: “explicit uncertainty” (l. 97), “analytical uncertainty” (l. 406), and, on Figure 3, “grid-based errors in the prior” and “global uncertainty”.

Response: We have not been consistent about the terminology, and particularly the use of terms such as uncertainty, error and variance. We have revised the manuscript so that we are consistent about the terminology, and have clarified what we mean by explicit uncertainty and analytical uncertainty. The changes made are listed in response to the earlier comment about the confusion between uncertainty and variance.

Comment: As the uncertainty quantification seems to be a selling point of the present article, the assessment should be more open and transparent about sources of uncertainty, and discuss which of theses sources can be quantified and how. For example, little is said about uncertainties introduced by the CO2 physiological correction. Is it guaranteed to be accurate?

Response: There are three basic sources of uncertainty: the pollen-based reconstructions, the construction of the prior, and the uncertainties associated with our implementation of the method. We addressed the uncertainties associated with the first two, but the methodological uncertainties were not as well addressed in this paper (although they are discussed in the Prentice et al., 2017 paper from which we derived the CO2 correction approach, and in the arXiv pre-print). The expanded description of the variational method (see above) is now more explicit about potential uncertainties associated e.g. with choice of length scales and cut-offs. For the CO2 correction, we made a series of sensitivity analyses in the Prentice et al. (2017) paper to determine the impact of uncertainties (or errors) in the input parameters. These sensitivity tests showed that the magnitude of the correction was insensitive to the re-constructed temperature, the reconstructed change in temperature relative to the modern reference, or the reconstructed moisture level. The magnitude of the correction is highly sensitive to the level of CO2 specified, but this is well-constrained from the ice-core records. We have expanded the text in the discussion of the CO2 effect to make this clearer (lines 378-385), as follows:

. . .. differences in water use efficiency of different PFTs can be almost entirely accounted for by a single equation, as proposed here. Sensitivity analyses show that the numerical value of the corrected moisture variables (MI, MAP) is dependent on the reconstructed values of these variables but is insensitive to uncertainties in the temperature and moisture inputs (Prentice et al., 2017). The strength of the correction is primarily sensitive to [CO2], but the LGM [CO2] value is well constrained from ice-core records. The response of plants to changes in [CO2] is non-linear (Harrison and Bartlein, 2012), and the effect of the change between recent and pre-industrial or mid-Holocene conditions is less than that between pre-industrial and glacial conditions. Nevertheless, it would be worth taking the [CO2] effect on water-use efficiency into account in making reconstructions of interglacial time periods as well.

Comment: The strategy for identifying grid points with little posterior update explained l. 406 is not quite clear. Why not consider a Kullback-Leibler divergence? At the risk of repeating myself, I am concerned about the (meaningless) residual influence of the prior variance and covariance in cases where the prior is effectively discarded by the observations.

Response: As we have explained above in response to the question about masking (and will clarify in the text, lines 277-278), for each variable in each grid cell, we calculate the percentage change of variance between the prior and posterior. We then mask away variables where there is a less than 5% increase in variance. We do not use the Kullback-Leibler divergence approach because this requires the calculation of covariance. However, the two approaches will likely not yield results that are very dissimilar.

Comment. … the comparison with Goosse et al. 2006 is perhaps slightly misleading. The Goosse et al. purpose was dynamic reconstruction, while the purpose of the present contribution is to provide a diagnostic reconstruction. In passing, Goosse (2006) did not use a "Kalman particle filter" (whatever it means). Goosse et al. used what they called an "optimal realisation" iteration, which can be interpretated as a highly degenerate form of particle filter. Dubinkina et al. 2011, doi 10.1142/S0218127411030763, adopted a more standard particle filter.

Response: The reference to the Kalman filter is somewhat misleading, although the approach used by Goosse et al. (2006) can be considered a form of particle filter. Our point here is that filters that select from model output are inherently constrained by the model output, whereas variational approaches can go beyond the values produced by the model. We have changed the wording of the text (line 420-422) to make this clearer:

Particle filter approaches (e.g. Goosse et al., 2006; Dubinkina et al., 2011) produce dynamic estimates of palaeoclimate, but particle filters cannot produce estimates of climate outside the realm of the model simulations.

Comment: This said, the argument that the variational approach produces maps outside the realm of climate simulations is a double-edged sword. The variational approach assumes Gaussian distributions, and is mathematically equivalent to a Laplace approximation of arbitrary distributions. This is this approximation which allows generating posterior distributions far from the prior. But, in this case, sound Bayesian interpretation should lead us to treat such posterior as utterly suspicious.

Response: It is not clear why a posterior distribution that is far from the model-based prior is utterly suspicious, if being far from the prior reflects the fact that the observational constraints are strong. We are not pretending that there should be equal weight given to the model-based prior and the pollen-based reconstructions, only combining the two and drawing on their individual strengths produces a more reliable estimate of the "true" climate state. Our approach is specifically designed to permit analytical reconstructions that are far from the model-based prior, if this is consistent with the observations and those observations have low variance.

Comment: line 384 : It is said that it "would be worth taking [changes in length scales] into account." I would advise either deleting this sentence, or giving more substance to the claim. For example, have you already performed some sensitivity experiments.

Response: The cited text is not talking about changes in length scales, but rather about the necessity to take the CO2 correction into account in making reconstructions of interglacial climates. We have amended this sentence to make this clear, as follows:

Nevertheless, it would be worth taking the [CO2] effect on water-use efficiency into account in making reconstructions of interglacial time periods as well.

Comment: Is the very first paragraph really necessary?

Response: Strictly speaking, it should not be necessary, especially for a palaeoclimate audience. However, this does seem to be a point which is largely ignored by many climate modelling centres worldwide, and is therefore worth repeating.

Comment: There is room for improving wording accuracy. In what sense is the benchmark "ro- bust" (l. 37) ? l. 97: You write: "explicit uncertainties attached to it". Did you mean "uncertainties explicitly attached" ? Avoid, where possible, the phrase "in terms of" or "means that" (ll. 321 - 326, in particular, need rewording). What is meant by a "statistical reconstruction method" l. 370 (the present exercise is a statistical reconstruction isnt'it ?).

Response: We have been through the manuscript and tightened up the wording. With respect to the specific sentences above, we have made the following changes:

L 37: Thus, the new reconstructions provides a benchmark created using clear and defined mathematical procedures that can be used for evaluation of the PMIP4/CMIP6 entry-card LGM simulations …

L. 97: However, there has so far been no attempt to produce a physically consistent, multi-variable reconstruction which provides the associated uncertainties explicitly.

L 321 et seq.: There are systematic differences, however, between the analytical re-constructions and the pollen-based reconstructions of moisture-related variables (MAP, MI) because the analytical reconstructions take account of the direct influence of [CO2] on plant growth. The physiological impact of [CO2] leads to analytical reconstructions indicating wetter than present conditions in many regions (Figure 5a, Figure 5b), for example in southern Africa where several of the original pollen-based reconstructions show no change in MAP or MI compared to present, but the analytical reconstruction shows wetter conditions than present. In some regions, incorporating the impact of [CO2] reverses the sign of the reconstructed changes. Part of northern Eurasia is reconstructed as being wetter than present, despite pollen-based reconstructions indicating conditions drier than present (both in terms of MAP and MI), as shown by SI Figure 3. The relative changes in MAP and MI are similar across all sites (Figure 5c), implying that the analytically reconstructed changes are driven by changes in precipitation rather than temperature.

L 370: Statistical reconstruction methods that use modern relationships between pollen assemblages and climate under modern conditions (i.e. modern analogues, transfer functions, response surfaces: see Bartlein et al., 2011 for discussion) cannot account for such effects.

Comment: Figure 5: Shouldn't "pre-industrial reference" be preferred over the vague wording "original" as x-axis label?

Response: The axis labels on this Figure are not clear. These plots contrast the original pollen-based reconstructions of MI and MAP with analytical re-constructions before (circles) and after (crosses) the CO2 effect is taken into account. We have changed the axis labels to read: Pollen-based MI and Pollen-based MAP. We have expanded the caption to make this clearer, as follows:

Figure 5: Impact of CO_2 on reconstructions of moisture-related variables. The individual plots show (a) the change in moisture index (MI) and (b) the change in mean annual precipitation (MAP) compared to the original pollen-based reconstructions for the LGM when the physiological impacts of [CO2] on water-use efficiency are taken into account. The third plot (c) shows the relative difference in MI and MAP as a result of [CO2], shown as the percentage difference between the no-[CO2] and [CO2] calculations.

Response to review 3

Comment: The paper is quite short and lacks any detailed evaluation of the resultant product. The. community's use of this new data product would in my opinion be aided by a more in-depth evaluation of the properties of the reconstruction. It's not clear how important thechoices around the assimilation formulation are for the final reconstruction. Specifically the section around lines 268-278 should in my opinion be spelled out and the sensitivity to these choices evaluated.

Response: It is unclear what kind of evaluation of the product the reviewer envisages, given that there is no global ground-truth data set other than the pollen-based reconstructions themselves. We have already pointed out (lines 317-321) that the analytical reconstructions of temperature are close to the Bartlein et al. (2011) data set, both in terms of magnitudes and spatial patterns. The differences between the analytical reconstructions and the Bartlein et al. (2011) reconstructions of moisture variables are a consequence of the fact that statistical techniques based on modern pollen-climate relationships cannot account for CO2-induced changes in water-use efficiency. In terms of the impact of methodological choices, the major issue here is the choice of length scales. We have made sensitivity analyses to examine the implications of the choice of length scales, and this was discussed in the arXiv preprint. In expanding the description of the application of variational techniques here (see text in response to Michel Crucifix's review) we have commented further on this.

Commeent: The statistical methodology that forms the basis of this study is also not described here but in a arXiv article. I'd like to see more of this brought into the present manuscript to make it self-contained.

Response: This is a point raised by Michel Crucifix in his review. We have now modified the text describing the application of the variational method to include a fuller description of our approach. The full details of the method are now in press in JAMES and we have made the post-review version of this paper available on arXiv.

Comment: Line 127: define MI here.

Response: The reference to MI is inappropriate in the present context because the text refers to a generic control by moisture availability rather than a specific index. We also note there was a crucial comma missing in this sentence! In response to Michel Crucifix's review, we have modified this text to read:

which is generally taken into account by process-based ecosystem models, but not by statistical models, using projected changes in vapour pressure deficit or some measure of plant-available water

Comment: Line 209: modelsfor -> models for
Response: We have corrected this.

Comment: Line 252-253: I think it might be appropriate to bring some/all of this methodology into the present text, as discussed above.

Response: We have modified the text here to provide more detail about the method. Please see proposed revised text given in the response to Michel Crucifix's review.

Comment: One question that arises from briefly reading the methodology paper, relates to figure 1 in the arXiv article. Here the assimilation appears not satisfy the pollen-inferred MTCO. Is this because the prior (from the models) is relatively consistent, and so doesn't allow the assimilation to get that cold? Does this happen when applied to the pollen data here?

Response: Figure 1 in the arXiv pre-print does not show a real situation but was designed to illustrate the procedure. In general, the pollen-based reconstructions of MTCO are further away from the model-based prior then summer temperature measures. If the variance in the pollen-based MTCO reconstructions. is small, then the analytical reconstructions will be close to the pollen-inferred MTCO. If there is high uncertainty in the pollen-based reconstructions, then the analytical reconstructions are not strongly constrained by these reconstructions and will be further away. This makes intuitive sense because we do not want to rely on pollen-based reconstructions if there is large uncertainty. Thus, it is possible for the assimilation to produce cold results but only if there is tight agreement between the observations about the magnitude of the cooling.

Comment: How do we interpret these choices, given that the climate models themselves could feasibly be systematically biased, e.g. through not including aerosols, or using modern vegetation distributions? How have you addressed the possible systematic bias in the models and hence in your prior?

Response: It is possible that the models show a systematic bias because they do not include all of the appropriate forcings for the LGM climate. We assume that such a systematic bias would primarily influence the magnitude of changes rather than the physical relationship between variables or across space. The presence of a systematic bias is therefore not important because the pollen-based reconstructions effectively correct for any systematic biases in the model-based prior, providing the pollen-based reconstructions have low uncertainty. One of the reasons that we discuss in the paper for adopting a variational technique, rather than some kind of filtering, is that this approach means that the analytical reconstructions can go beyond the range of the simulated climate.

Comment: Line 268-276: This section seems crucial to me, but is not clearly described. Please include the mathematical formulation used and a justification for choices made.

Response: We have expanded the text describing the application of the variational method, including a description of the composite errors. Please see revised text included in the response to Michel Crucifix's review.

Comment Lines 276-278: Do you mean that if the data is too uncertain you mask it based on a 5% criteria? Please could you re-phrase to clarify.

Response: When the change in the variance between the analytical reconstruction and the prior is less than 5%, it does indeed mean that the climate is not well constrained by observations (i.e. that there is high uncertainty in the observations). We have modified this text (and the discussion of the choice of cutoff in the Discussion) to clarify this point. Please see revised text in response to Michel Crucifix's review.

Comment Lines 288: How does your product compare with the original Bartlein et al 2011, and the GCM-based prior? Could you show this?

Response: The GCM-based prior is shown in SI Figure 1 and the original pollen-based reconstructions (from Bartlein et al, 2011 and from Prentice et al., 2017) are shown in SI Figure 3. Comparison of these figures with the analytical reconstructions shown in the paper in Figure 4 shows the difference with our product. We could add a new set of figures to the Supplementary showing difference maps, if necessary.

Comment: How well is the seasonality captured and how does it differ from the simulated seasonality in the GCM prior?

Response: We have no independent measure of seasonality that can be used to assess the analytical reconstructions. The analytical reconstructions of MTCO and MTWA, the difference between which is the measure of the strength of temperature seasonality, are only shown when the pollen-based reconstructions contain sufficient information to modify the model-based prior and thus when the uncertainty in the pollen-based reconstructions is small. We could produce maps showing the temperature seasonality from the analytical reconstructions and the model ensemble (and the difference between them) but these would not add anything to the manuscript beyond what is shown by the MTCO and MTWA reconstructions.

Editor's Comments

Comment: line 272: the "-" in "-5.6" should be a real minus sign.
Response: We have changed this.

Comment: Figure 4 shows climate anomalies w.r.t the CRU average for 1960-1990. This should be clearly stated in the caption.
Response:

Comment: line 859: the "e" should be in italic
Response: We have changed this

We have added a code/data availability section, author contributions and competing interests statement.

A new multi-variable benchmark for Last Glacial Maximum climate simulations

Sean F. Cleator[1], Sandy P. Harrison[2], Nancy K. Nichols[3], I. Colin Prentice[4] and Ian Roulstone[1]

1: Department of Mathematics, University of Surrey, Guildford GU2 7XH, UK
2: School of Archaeology, Geography and Environmental Science, University of Reading, Whiteknights, Reading, RG6 6AH, UK
3: Department of Mathematics & Statistics, University of Reading, Whiteknights, Reading RG6 6AX, UK
4: AXA Chair of Biosphere and Climate Impacts, Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot SL5 7PY, UK

Journal: Climate of the Past

Abstract. We present a new global reconstruction of seasonal climates at the Last Glacial Maximum (LGM, 21,000 yr BP) made using 3-D variational data assimilation with pollen-based site reconstructions of six climate variables and the ensemble average of the PMIP3/CMIP5 simulations as a prior. We assume that the correlation matrix of the uncertainties of the prior both spatially and temporally is Gaussian, in order to produce a climate reconstruction that is smoothed both from month to month and from grid cell to grid cell. The pollen-based reconstructions include mean annual temperature (MAT), mean temperature of the coldest month (MTCO), mean temperature of the warmest month (MTWA), growing season warmth as measured by growing degree days above a baseline of 5°C (GDD_5), mean annual precipitation (MAP) and a moisture index (MI), which is the ratio of MAP to mean annual potential evapotranspiration. Different variables are reconstructed at different sites, but our approach both preserves seasonal relationships and allows a more complete set of seasonal climate variables to be derived at each location. We further account for the ecophysiological effects of low atmospheric carbon dioxide concentration on vegetation in making reconstructions of MAP and MI. This adjustment results in the reconstruction of wetter climates than might otherwise be inferred by the vegetation composition. Finally, by comparing the uncertainty contribution to the final reconstruction, we provide confidence intervals on these reconstructions and delimit geographical regions for which the palaeodata provide no information to constrain the climate reconstructions. The new reconstructions will provide a benchmark created using clear and defined mathematical procedures that can be used for evaluation of the PMIP4/CMIP6 entry-card LGM simulations and are available at DOI:10.17864/1947.229.

Deleted: errors

Deleted: error

Deleted: robust

1 Introduction

Models that perform equally well for present-day climate nevertheless produce very different responses to anthropogenic forcing scenarios through the 21st century. Although internal variability contributes to these differences, the largest source of uncertainty in model projections in the first three to four decades of the 21st century stems from differences in the response of individual models to the same forcing (Kirtman et al., 2013). Thus, the evaluation of models based on modern observations is not a good guide to their future performance, largely because the observations used to assess model performance for present-day climate encompass too limited a range of climate variability to provide a robust test of the ability to simulate climate changes. Although past climate states do not provide analogues for the future, past climate changes provide a unique opportunity for out-of-sample evaluation of climate model performance (Harrison et al., 2015).

At the Last Glacial Maximum (LGM, conventionally defined for modelling purposes as 21 000 years ago), insolation was quite similar to the present, but global ice volume was at a maximum, eustatic sea level was close to a minimum, long-lived greenhouse gas concentrations were lower, and atmospheric aerosol loadings higher than today, and land surface characteristics (including vegetation distribution) were also substantially different from today. These changes gave rise to a climate radically different from that of today; indeed the magnitude of the change in radiative forcing between LGM and pre-industrial climate is comparable to high-emissions projections of climate change between now and the end of the 21st century (Braconnot et al., 2012). The LGM has been a focus for model evaluation in the Paleoclimate Modelling Intercomparison Project (PMIP) since its inception (Joussaume and Taylor, 1995; Braconnot et al., 2007; Braconnot et al., 2012). The LGM is one of the two “entry card” palaeoclimate simulations included in the current phase of the Coupled Model Intercomparison Project (CMIP6) (Kageyama et al., 2018). The evaluation of previous generations of palaeoclimate simulations has shown that the large-scale thermodynamic responses seen in 21st century and LGM climates, including enhanced land–sea temperature contrast, latitudinal amplification, and scaling of precipitation with temperature, are likely to be realistic (Izumi et al., 2013; Li et al., 2013; Lunt et al, 2013; Hill et al., 2014; Izumi et al., 2014; Harrison et al., 2015). However, evaluation against palaeodata shows that even when the sign of large-scale climate changes is correctly predicted, the patterns of change at a regional scale are often inaccurate and the magnitudes of change often underestimated (Brewer et al., 2007; Mauri et al., 2014; Perez Sanz et al., 2014; Bartlein et al., 2017). The current focus on understanding what causes mismatches between reconstructed and simulated climates is a primary motivation for developing benchmark data sets that represent regional climate changes comprehensively enough to allow a critical evaluation of model deficiencies.

Many sources of information can be used to reconstruct past climates. Pollen-based reconstructions are the most widespread, and pollen-based data were the basis for the current standard LGM benchmark data set by Bartlein et al. (2011). In common with other data sources, the pollen-based reconstructions were generated for individual sites. Geological preservation issues mean that the number of sites available inevitably decreases through time (Bradley, 2014). Since pollen is only preserved for a long time in anoxic sediments, the geographic distribution of potential sites is biased towards climates that are relatively wet today. Furthermore, the actual sampling of potential sites is highly non-uniform, so there are large geographic gaps in data coverage (Harrison et al., 2016). The lack of continuous climate fields is not ideal for model evaluation, and so attempts have been made to generalize the site-based data either through gridding, interpolation, or some form of multiple regression (see e.g. Bartlein et al., 2011; Annan and Hargreaves, 2013). However, there has so far been no attempt to produce a physically consistent, multi-variable reconstruction which provides the associated uncertainties explicitly.

Deleted: However, there has so far been no attempt to produce a physically consistent, multi-variable reconstruction with explicit uncertainties attached to it.

A further characteristic of the LGM that creates problems for quantitative reconstructions based on pollen data is the much lower atmospheric carbon dioxide concentration, [CO_2], compared to the pre-industrial Holocene. [CO_2] has a direct effect on plant physiological processes. Low [CO_2] as experienced by plants at the LGM is expected to have led to reduced water-use efficiency – the ratio of carbon assimilation to the water lost through transpiration (Bramley et al., 2013). Most reconstructions of moisture variables from pollen data, including most of the reconstructions used by Bartlein et al. (2011), do not take [CO_2] effects into account. Yet several modelling studies have shown that the impact of low [CO_2] around the LGM on plant growth and distribution was large (e.g. Jolly and Haxeltine, 1997; Cowling and Sykes, 1999; Harrison and Prentice, 2003; Bragg et al., 2013; Martin Calvo et al., 2014; Martin Calvo and Prentice, 2015). A few reconstructions of LGM climate based on the inversion of process-based biogeography models have also shown large effects of low [CO_2] on reconstructed LGM palaeoclimates (e.g. Guiot et al., 2000; Wu et al., 2007). The reconstructions of moisture variables in the Bartlein et al. (2011) data set are thus probably not reliable, and likely to be biased low.

Prentice et al. (2017) demonstrated an approach to correct reconstructions of moisture variables for the effect of [CO_2], but this correction has not been applied globally. A key side effect of applying this [CO_2] correction is to reconcile semi-quantitative hydrological evidence for wet conditions at the LGM with the apparent dryness suggested by the vegetation assemblages (Prentice et al., 2017). Similar considerations apply to the interpretation of future climate changes in terms of vegetational effects. Projections of future aridity (based on declining indices of moisture availability) linked to warming are unrealistic, in a global perspective, because of the counteracting effect of increased water use efficiency due to rising [CO_2] – which is generally taken into account by process-based ecosystem models, but not by statistical models, using projected changes in vapour pressure deficit or some measure of plant-available water (Keenan et al., 2011; Roderick et al., 2015; Greve et al., 2017).

Deleted: relying on

Deleted: or MI

In this paper, we use variational data assimilation based on both pollen-based climate reconstructions and climate model outputs to arrive at a best-estimate analytical reconstruction of LGM climate, explicitly taking account of the impact of [CO_2]. Variational techniques provide a way of combining observations and model outputs to produce climate reconstructions that are not exclusively constrained to one source of information or the other (Nichols, 2010). We use the uncertainty contributions to the analytical reconstruction to provide confidence intervals for these reconstructions and also to delimit geographical regions for which the palaeodata provide no constraint on the reconstructions. The resulting data set is expected provide a well-founded multi-variable LGM climate dataset for palaeoclimate model benchmarking in CMIP6.

Deleted: error

2 Methods

2.1 Pollen-based climate reconstructions

Bartlein et al. (2011) provided a global synthesis of pollen-based quantitative climate reconstructions for the LGM. The Bartlein et al. (2011) data set includes reconstructions of climate anomalies (differences between LGM and recent climates) for six variables (and their uncertainties), specifically mean annual temperature (MAT), mean temperature of the coldest month (MTCO), mean temperature of the warmest month (MTWA), growing degree days above a baseline of above 5°C (GDD5), mean annual precipitation (MAP), and an index of plant-available moisture (the ratio of actual to equilibrium evapotranspiration, or α). There are a small number of LGM sites (94) in the Bartlein et al. (2011) data set where model inversion was used to make the reconstructions of α and MAP; no [CO_2] correction is applied to these reconstructions. There are no data from Australia in the Bartlein et al. (2011) data set, and we therefore use quantitative reconstructions of MAT and another moisture index (MI), the ratio of MAP to potential evapotranspiration, from Prentice et al. (2017). Prentice et al. (2017) provide values of MI both before and after correction for [CO_2]; we use the uncorrected values in order to apply the correction for [CO_2] within our assimilation framework. For consistency between the two data sets, we re-expressed reconstructions of α in terms of MI via the Fu-Zhang formulation of the Budyko relationship between actual evapotranspiration, potential evapotranspiration and precipitation (Zhang et al., 2004; Gallego-Sala et al., 2016).

Deleted: these were excluded from our analysis

The spatial coverage of the final data set is uneven (Figure 1). There are many more data points in Europe and North America than elsewhere. South America has the fewest

(14 sites). The number of variables available at each site varies: although most sites (279) have reconstructions of at least three variables, some sites have reconstructions of only one variable (60). Nevertheless, in regions where there is adequate coverage, the reconstructed anomaly patterns are coherent, plausible and consistent among variables.

For this application, we derived absolute LGM climate reconstructions by adding the reconstructed climate anomalies at each site to the modern climate values from the Climate Research Unit (CRU) historical climatology data set (CRU CL v2.0 dataset, New et al., 2002), which provides climatological averages of monthly temperature, precipitation and cloud cover fraction for the period 1961-1990 CE. Most of the climate variables (MTCO, MTWA, MAT, MAP) can be calculated directly from the CRU CL v2.0 dataset. GDD5 was calculated from pseudo-daily data derived by linear interpolation of the monthly temperatures. MI was calculated from the CRU climate variables using the radiation calculations in the SPLASH model (Davis et al., 2017). For numerical efficiency, we non-dimensionalised all of the absolute climate reconstructions (and their standard errors) before applying the variational techniques (for details, see Cleator et al., 2019a).

2.2 Climate model simulations

Eight LGM climate simulations (Table 1) from the third phase of the Palaeoclimate Modelling Intercomparison Project (PMIP3: Braconnot et al., 2012) were used to create a prior. The PMIP LGM simulations were forced by known changes in incoming solar radiation, changes in land-sea geography and the extent and location of ice sheets, and a reduction in [CO_2] to 185 ppm (see Braconnot et al., 2012 for details of the modelling protocol). We used the last 100 years of each LGM simulation. We interpolated monthly precipitation, monthly temperature and monthly fraction of sunshine hours from each LGM simulation and its pre-industrial (PI) control to a common 2 x 2° grid. Simulated climate anomalies (LGM minus PI) for each grid cell were then added to modern climate values calculated from the CRU CL 2.0 data set (New et al., 2002), as described for the pollen-based reconstructions, to derive absolute climate values. We calculated the multi-model mean and variance (Figure 2) across the models for each of the climate variables to produce the gridded map used as the prior.

2.3 Water-use efficiency calculations

We applied the general approach developed by Prentice et al. (2017) to correct pollen-based statistical reconstructions to account for [CO_2] effects. The approach as implemented here is based on equations (Appendix 1) that link moisture index (MI) to transpiration and the ratio of leaf-internal to ambient CO_2. The correction is based on the principle that the rate of water loss per unit carbon gain is inversely related to effective moisture availability as sensed by plants. The method involves solving a non-linear equation that relates rate of water loss per unit carbon gain to MI, temperature and CO_2 concentration. The equation is derived from theory that predicts the response of the ratio of leaf-internal to ambient [CO_2] to vapour pressure deficit and temperature (Prentice et al., 2014; Wang et al., 2014).

2.4 Application of variational techniques

Variational data assimilation techniques provide a way of combining observations and model outputs to produce climate reconstructions that are not exclusively constrained to one source of information or the other (Nichols, 2010). We use the 3D-variational method, described in Cleator et al. (2019a) to find the maximum a posteriori estimate (or analytical reconstruction) of the palaeoclimate given the site-based reconstructions and the model-based prior. The method constructs a cost function, which describes how well a particular climate matches both the site-based reconstructions and the prior, by assuming the reconstructions and prior have a Gaussian distribution. To avoid sharp changes in time and/or space in the analytical reconstructions, the method assumes that the prior temporal and spatial covariance correlations are derived from a modified Bessel function, in order to create a climate anomaly field that is smooth both from month to month and from grid cell to grid cell. The degree of correlation is controlled through two length scales: a spatial length scale that determines how correlated the covariance in the prior is between different geographical areas, and a temporal length scale that determines how correlated it is through the seasonal cycle. The site-based reconstructions are assumed to have negligible correlations at these space and time scales. The maximum a posteriori estimate is found by using the limited memory Broyden- Fletcher-Goldfarb-Shanno method (Liu and Nocedal 1989) to determine the climate that minimises the cost function. A first order estimate of the analysis uncertainty covariance is also computed.

An observation operator based on calculations of the direct impact of [CO2] on water-use efficiency (section 2.3) is used in making the analytical reconstructions. The prior is constructed as the average of eight LGM climate simulations (section 2.2). We use an ensemble of different model responses to the same forcing to provide a series of physically consistent possible states, which can be viewed as perturbed responses and provide the variance around the climatology provided by the ensemble average. The prior uncertainty correlations are based on a temporal length scale (Lt) of 1 month and a spatial length scale (Ls) of 400km. Cleator et al., (2019a) have shown that a temporal length scale of 1 month provides an adequately smooth solution for the seasonal cycle, both using single sites and over multiple grid cells, as shown by the sensitivity of the resolution matrix (Menke, 2012; Delahaies et al., 2017) to changes in the temporal length scale. Consideration of the spatial spread of variance in the analytical reconstruction shows that a spatial length scale of 400km also provides a reasonable reflection of the large-scale coherence of regional climate change.

We generated composite variances on the analytical reconstructions (Figure 3) by combining the covariances from the site-based reconstructions and from the prior. There are regions where all of the models systematically differ from the site-based reconstructions (Harrison et al., 2015) but nevertheless the inter-model variability is low, which would lead to a very small contribution to the composite uncertainties from the prior. We therefore calculated the uncertainty of the prior from an equal combination of the global uncertainty, the average variance between each grid cell, and local uncertainty, the variance between the different models. The reliability of the analytical reconstructions was assessed by comparing these composite covariances with the uncertainties in the prior. We masked out cells where the inclusion of site-based reconstructions does not produce an improvement of > 5% from the prior. Since this assessment is based on a change in the variance, rather than absolute values, this masking removes regions where there are no pollen-based reconstructions or the pollen-based reconstructions have very large uncertainties.

3 Results

The analytical reconstructions (Figure 4) show an average year-round cooling of –5.6 °C in the northern extratropics. The cooling is larger in winter (–7.6 °C) than in summer (–2.4 °C). A limited number of grid cells in central Eurasia show warmer-than-present summers, and higher MAT. Temperature changes are more muted in the tropics, with an average change in MAT of –3.7 °C. The cooling is somewhat lower in summer than winter (–2.7 °C compared to –4.1 °C). Reconstructed temperature changes were slightly larger in the southern extratropics, with average changes in MAT of –5.0°C, largely driven by cooling in winter.

Deleted: -6.9
Deleted: 8.2
Deleted: 3.8
Deleted: 5
Deleted: 1
Deleted: 3.3
Deleted: smaller
Deleted: 0.8
Formatted: English (US)

Changes in moisture-related variables (MAP, MI) across the northern hemisphere are geographically more heterogeneous than temperature changes. Reconstructed MAP is greater than present in western North America (172 mm) but less than present (–29 mm) in eastern North America. Most of Europe is reconstructed as drier than present (–305mm), the same for eastern Eurasia (-94 mm) and the Far East (–66 mm). The patterns in MI are not identical to those in MAP, because of the influence of temperature

Deleted: 58
Deleted: 342
Deleted: 241
Deleted: 126
Deleted: 43

on MI, but regional changes are generally similar to those shown by MAP. Most of the tropics are shown as drier than present while the southern hemisphere extratropics are wetter than present, in terms of both MAP and MI.

The reconstructed temperature patterns are not fundamentally different from those shown by Bartlein et al. (2011) but the analytical dataset provides information for a much larger area (1153% increase) thanks to the method's imposition of consistency among different climate variables, and smooth variations both in space and through the seasonal cycle. There are systematic differences, however, between the analytical reconstructions and the pollen-based reconstructions in terms of moisture-related variables (MAP, MI) because the analytical reconstructions take account of the direct influence of [CO2] on plant growth. The physiological impact of [CO2] leads to analytical reconstructions indicating wetter than present conditions in many regions (Figure 5a, Figure 5b), for example in southern Africa where several of the original pollen-based reconstructions show no change in MAP or MI compared to present, but the analytical reconstruction shows wetter conditions than present. In some regions, incorporating the impact of [CO_2] reverses the sign of the reconstructed changes. Part of northern Eurasia is reconstructed as being wetter than present, despite pollen-based reconstructions indicating conditions drier than present (both in terms of MAP and MI), as shown by SI Figure 3. The relative changes in MAP and MI are similar across all sites (Figure 5c), implying that the analytically reconstructed changes are driven by changes in precipitation rather than temperature.

4 Discussion

Variational data assimilation techniques provide a way of combining observations and model outputs, taking account the uncertainties in both, to produce a best-estimate analytical reconstruction of LGM climate. These reconstructions extend the information available from site-based reconstructions both spatially and through the seasonal cycle. Our new analytical data set characterizes the seasonal cycle across a much larger region of the globe than the data set that is currently being used for benchmarking of palaeoclimate model simulations. We therefore suggest that this data set (Cleator et al. 2019b) should be used for evaluating the CMIP6-PMIP4 LGM simulations.

Some areas are still poorly covered by quantitative pollen-based reconstructions of LGM climate, most notably South America. More pollen-based climate reconstructions would provide one solution to this problem – and there are many pollen records that could be used for this purpose (Flantua et al., 2015; Herbert and Harrison, 2016; Harrison et al., 2016). There are also quantitative reconstructions of climate available from individual sites (e.g. Lebamba et al., 2012; Wang et al., 2014; Loomis et al., 2017;

Deleted: 643

Deleted: , and this largely reflects the influence of [CO_2] that is included in the analysis.

Deleted: Accounting for the physiological impact of [CO_2] means that the analytical reconstructions indicate wetter than present conditions in many regions

Deleted: SI

Deleted: (both in terms of MAP and MI)

Camuera et al., 2019) that should be incorporated into future data syntheses. It would also be possible to incorporate other sources of quantitative information, such as chironomid-based reconstructions (e.g. Chang et al., 2015), within the variational data assimilation framework.

One of the benefits of the analytical framework applied here is that it allows the influence of changes in [CO_2] on the moisture reconstructions to be taken into account. Low [CO_2] must have reduced plant water-use efficiency, because at low [CO_2] plants need to keep stomata open for longer in order to capture sufficient CO_2. Statistical reconstruction methods that use modern relationships between pollen assemblages and climate under modern conditions (i.e. modern analogues, transfer functions, response surfaces: see Bartlein et al., 2011 for discussion) cannot account for such effects. Climate reconstruction methods based on the inversion of process-based ecosystem models can do so (see e.g. Guiot et al., 2000; Wu et al., 2007; Wu et al., 2009; Izumi and Bartlein, 2016) but are critically dependent on the reliability of the vegetation model used. Most of the palaeoclimate reconstructions have been made by inverting some version of the BIOME model (Kaplan et al., 2003), which makes use of bioclimatic thresholds to separate different plant functional types (PFTs). As a result, reconstructions made by inversion show "jumps" linked to shifts between vegetation types dominated my different PFTs whereas, as has been shown recently (Wang et al., 2017), differences in water use efficiency of different PFTs can be almost entirely accounted for by a single equation, as proposed here. Sensitivity analyses show that the numerical value of the corrected moisture variables (MI, MAP) is dependent on the reconstructed values of these variables but is insensitive to uncertainties in the temperature and moisture inputs (Prentice et al., 2017). The strength of the correction is primarily sensitive to [CO2], but the LGM [CO2] value is well constrained from ice-core records. The response of plants to changes in [CO_2] is non-linear (Harrison and Bartlein, 2012), and the effect of the change between recent and pre-industrial or mid-Holocene conditions is less than that between pre-industrial and glacial conditions. Nevertheless, it would be worth taking the [CO2] effect on water-use efficiency into account in making reconstructions of interglacial time periods as well.

The influence of individual pollen-based reconstructions on the analytical reconstruction of seasonal variability, or the geographic area influenced by an individual site, is crucially dependent on the choice of length scales. We have adopted conservative length scales of 1 month and 400 km, based on sensitivity experiments made for southern Europe (Cleator et al., 2019a). These length scales produce numerically stable results for the LGM, and the paucity of data for many regions at the LGM means that using fixed, conservative length scales is likely to be the only practical approach. However, in so far as the spatial length scale is related to atmospheric circulation patterns, there is no reason to suppose that the optimal spatial length scale will be the same from region to region. The density and clustering of pollen-based

Deleted: Statistical reconstruction methods, whether based on modern analogues or modern climate transfer functions, cannot account for such effects.

Deleted: Nevertheless, it would be worth taking this effect into account in making reconstructions of interglacial time periods as well.

reconstructions could also have a substantial effect on the optimal spatial length scale. A fixed 1-month temporal length scale is appropriate for climates that have a reasonably smooth and well-defined seasonal cycle, either in temperature or precipitation. However, in climates where the seasonal cycle is less well defined, for example in the wet tropics, or in situations where there is considerable variability on sub-monthly time scales, other choices might be more appropriate. For time periods such as the mid-Holocene, which have an order of magnitude more site-based data, it could be useful to explore the possibilities of variable length scales.

Deleted:

We have used a 5% reduction in the analytical uncertainty compared to prior uncertainty to identify regions where the incorporation of site-based data has a negligible effect on the prior as a way of masking out regions for which the observations have effectively no impact on the analytical reconstructions. The choice of a 5% cut-off is arbitrary, but little would be gained by imposing a more stringent cut-off at the LGM given that many regions are represented by few observations. A more stringent cut-off could be applied for other time intervals with more data. We avoid the use of a criterion based on the analytical reconstruction showing any improvement on the prior because this could be affected by numerical noise in the computation. Alternative criteria for the choice of cut-off could be based on whether the analytical reconstruction had a reduced uncertainty compared to the pollen-based reconstructions or could be derived by a consideration of the condition number used to select appropriate length scales.

There have been a few previous attempts to use data assimilation techniques to generate spatially continuous palaeoclimate reconstructions. Annan and Hargreaves (2013) used a similar multi-model ensemble as the prior and the pollen-based reconstructions from Bartlein et al. (2011) to reconstruct MAT at the LGM. However, they made no attempt to reconstruct other seasonal variables, either independently, or through exploiting features of the simulations (as we have done here) to generate seasonal reconstructions. Particle filter approaches (e.g. Goosse et al., 2006; Dubinkina et al., 2011) produce dynamic estimates of palaeoclimate, but particle filters cannot produce estimates of climate outside the realm of the model simulations. Our 3-D variational data assimilation approach has the great merit of being able to produce seasonally coherent reconstructions generalized over space, while at the same time being capable of producing reconstructions that are outside those captured by the climate model, because they are not constrained by a specific source (Nichols, 2010). This property is of particular importance if the resulting data set is to be used for climate model evaluation, as we propose.

Deleted: Kalman particle filter approaches (e.g. Goosse et al., 2006) produce seasonal and geographical estimates of palaeoclimate, but particle filters cannot produce estimates of climate outside the realm of the model simulations.

Data availability. The gridded data for the LGM reconstructions are available from DOI:10.17864/1947.229; the code used to generate these reconstructions is available from DOI:10.5281/zenodo.3445166.

Author Contributions

All authors contributed to the design of the study; ICP developed the theory underlying the CO_2 correction; SC implemented the analyses. SC and SPH wrote the first version of the manuscript, and all authors contributed to the final version.

Competing Interests.

The authors declare they have no competing interests.

Acknowledgements

SFC was supported by a UK Natural Environment Research Programme (NERC) scholarship as part of the SCENARIO Doctoral Training Partnership. SPH acknowledges support from the ERC-funded project GC 2.0 (Global Change 2.0: Unlocking the past for a clearer future, grant number 694481). ICP acknowledges support from the ERC under the European Union's Horizon 2020 research and innovation programme (grant agreement no: 787203 REALM) This research is a contribution to the AXA Chair Programme in Biosphere and Climate Impacts and the Imperial College initiative on Grand Challenges in Ecosystems and the Environment (ICP). NKN is supported in part by the NERC National Center for Earth Observation (NCEO). We thank PMIP colleagues who contributed to the production of the palaeoclimate reconstructions. We also acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and the climate modelling groups in the Paleoclimate Modelling Intercomparison Project (PMIP) for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. The analyses and figures are based on data archived at CMIP on 12/09/18.

References

Annan, J.D., and Hargreaves, J.C.: A new global reconstruction of temperature changes at the Last Glacial Maximum, Clim. Past, 9, 367-376, https://doi.org/10.5194/cp-9-367-2013, 2013.

Bartlein, P.J., Harrison, S.P., Brewer, S., Connor, S., Davis, B.A.S., Gajewski, K., Guiot, J., Harrison-Prentice, T.I., Henderson, A., Peyron, O., Prentice, I.C., Scholze, M., Seppa, H., Shuman, B., Sugita, S., Thompson, R.S., Viau, A.E., Williams, J., and Wu, H.: Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis, Clim. Dynam., 37, 775-802, https://doi.org/10.1007/s00382-010-0904-1 2011.

Bartlein, P.J., Harrison, S.P., and Izumi, K.: Underlying causes of Eurasian mid-continental aridity in simulations of mid-Holocene climate, Geophys. Res. Letters, 44, doi: 10.1002/2017GL074476, 2017.

Braconnot, P., Otto-Bliesner, B, Harrison, S.P., Joussaume, S., Peterschmitt, J-Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, Th., Hewitt, C.D., Kageyama, M., Kitoh, A., Loutre, M-F., Marti, O., Merkel, U., Ramstein, G., Valdes, P., Weber, L., Yu, Y., and Zhao, Y.: Results of PMIP2 coupled simulations of the mid-Holocene and Last Glacial Maximum, Part 1: experiments and large-scale features, Clim. Past, 3, 261-277, doi:10.5194/cp-3-261-2007, 2007.

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, Nature Clim. Change, 2, 417-424, https://doi.org/10.1038/nclimate1456, 2012.

Bradley RS. 2014. Paleoclimatology: Reconstructing Climates of the Quaternary, 3rd edn. Academic Press/Elsevier: Amsterdam.

Bragg, F., Prentice, I.C., Harrison, S.P., Foster, P.N., Eglinton, G., Rommerskirchen F., and Rullkötter, J.: n-Alkane stable isotope evidence for CO_2 as a driver of vegetation change, Biogeosci., 10, 2001–2010, https://doi.org/10.5194/bg-10-2001-2013, 2013.

Bramley, H., Turner, N., and Siddique, K.: Water use efficiency. In Kole, C. (Ed.), *Genomics and Breeding for Climate-Resilient Crops* (1 ed., Vol. 2, pp. 487). Heidelberg: Springer. https://doi.org/10.1007/978-3-642-37048-9, 2013.

Brewer, S., Guiot, J., and Torre, F.: Mid-Holocene climate change in Europe: A data–model comparison, Clim. Past*,* 3**,** 499–512, https://doi.org/10.5194/cp-3-499-2007, 2007.

Budich, R., Giorgetta, M., Jungclaus, J., Redler, R., and Reick, C.: The MPI-M Millennium Earth System Model: An Assembling Guide for the COSMOS Configuration. https://pure.mpg.de/rest/items/item_2193290_2/component/file_2193291/content, 2010 (last accessed 31/03/2019).

Camuera, J., Jimenez-Moreno, G., Ramos-Roman, M.J., Garcia-Alix, A., Toney, J.L., Anderson, R.S., Jimenez-Espejo, F., Bright, J., Webster, C., Yanes, Y., Carrion, J.S.: Vegetation and climate changes during the last two glacial-interglacial cycles in the western Mediterranean: A new long pollen record from Padul (southern Iberian Peninsula). Quat. Sci. Rev., 205, 86-105, https://doi.org/10.1016/j.quascirev.2018.12.013, 2019.

Chang, J.C., Shulmeister, J., Woodward, C., Steinberger, L., Tibby, J., Barr, C.: A chironomid-inferred summer temperature reconstruction from subtropical Australia during the last glacial maximum (LGM) and the last deglaciation. Quat. Sci. Rev., 122, 282- 292, https://doi.org/10.1016/j.quascirev.2015.06.006, 2015.

Cleator, S.F., Harrison, S.P., Nichols, N.K., Prentice, I.C., and Roustone, I.: A method for generating coherent spatially explicit maps of seasonal palaeoclimates from site-based reconstructions, arXiv:1902.04973 [math.NA], 2019a.

Cleator, S., Harrison, S. P., Nichols, N., Prentice, I. C., Roulstone, I.: A new multi-variable benchmark for Last Glacial Maximum climate simulations. University of Reading. Dataset. DOI:10.17864/1947.206, 2019b,

Cowling, S.A., and Sykes, M.T.: Physiological significance of low atmospheric CO_2 for plant-climate interactions, Quatern. Res., 52, 237–242, DOI: 10.1006/qres.1999.2065, 1999.

Davis, T.W., Prentice, I.C., Stocker, B.D., Thomas, R.T., Whitley, R.J., Wang, H., Evans, B.J., Gallego-Sala, A.V., Sykes, M.T., and Cramer, W.: Simple process-led algorithms for simulating habitats (SPLASH v.1.0): robust indices of radiation, evapotranspiration and plant-available moisture, Geosci. Model Develop., 10, 689-708, https://doi.org/10.5194/gmd-10-689-2017, 2017.

Delahaies, S., Roulstone, I., Nichols, N.: Constraining DALECv2 using multiple data streams and ecological constraints: analysis and application. Geosci. Model Develop., 10, doi: 10.5194/gmd-10-2635- 2017, 2017.

Dubinkina, S., Goosse, H., Sallaz-Damaz, Y., Crespin, E., Crucifix, M.: Testing a particle filter to reconstruct climate changes over the past centuries. Int. J. Bifurcation Chaos 21, 3611-3618, 2011.

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, 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., Viovy, N., Vuichard, N.: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5, Clim. Dyn., 40, 2123-2165, DOI: 10.1007/s00382-012-1636-1, 2013.

Flantua, S.G.A., Hooghiemstra, H., Grimm, E.C., Behling, H., Bush, M.B., Gonzalez-Arango, C., Gosling, W.D., Ledru, M.P., Lozano-Garcia, S., Maldonado, A., Prieto, A.R., Rull, V., and Van Boxel, J. H.: Updated site compilation of the Latin American Pollen Database, Rev. Palaeobot. Palyno., 223, 104–115, https://doi.org/10.1016/j.revpalbo.2015.09.008, 2015.

Gallego-Sala, A., Charman, D., Li, G., Harrison, S.P., and Prentice, I.C.: Climate driven expansion of blanket bogs in the British Isles during the Holocene, Clim. Past, 12, 129-136, https://doi.org/10.5194/cp-12-129-2016, 2016.

Gent, P.R., Danabasoglu, G., Donner, L.J., Holland, M.M., Hunke, E.C., Jayne, S.R., Lawrence, D.M., Neale, R.B., Rasch, P.J., Vertenstein, M., Worley, P.H., Yang, Z-L., and Zhang, M.: The community climate system model version 4, J. Clim., 24, 4973-4991, https://doi.org/10.1175/2011JCLI4083.1, 2011.

Goosse, H., Renssen, H., Timmermann, A., Bradley, R.S., and Mann, M.E.: Using palaeoclimate proxy-data to select optimal realisations in an ensemble of simulations of the climate of the past millennium, Clim. Dynam., 27, 165-184, DOI 10.1007/s00382-006-0128-6, 2006.

Greve, P., Roderick, M.L., and Seneviratne, S.I.: Simulated changes in aridity from the last glacial maximum to $4xCO_2$. Environ. Res. Letters, 12, 114021, https://iopscience.iop.org/article/10.1088/1748-9326/aa89a3/meta, 2017.

Guiot, J., Torre, F., Jolly, D., Peyron, O., Boreux, J.J., and Cheddadi, R.: Inverse vegetation modeling by Monte Carlo sampling to reconstruct palaeoclimates under changed precipitation seasonality and CO_2 conditions: application to glacial climate in Mediterranean region. Ecological Modelling 127: 119–140, DOI: 10.1016/S0304-3800(99)00219-7, 2000.

Harrison, S.P., and Bartlein, P.J.: Records from the past, lessons for the future: what the palaeo-record implies about mechanisms of global change, in: Henderson-Sellers, A., and McGuffie, K. (Eds.), The Future of the World's Climate. Elsevier, pp. 403-436, 2012.

Harrison, S.P., and Prentice, I.C.: Climate and CO_2 controls on global vegetation distribution at the last glacial maximum: analysis based on palaeovegetation data, biome modelling and palaeoclimate simulations, Glob. Change Biol., 9, 983-1004, DOI: 10.1046/j.1365-2486.2003.00640.x, 2003.

Harrison, S.P., Bartlein, P.J., Brewer, S., Prentice, I.C., Boyd, M., Hessler, I., Holmgren, K., Izumi, K., and Willis, K.: Climate model benchmarking with glacial and mid-Holocene climates, Clim. Dynam., 43, 671-688, DOI: 10.1007/s00382-013-1922-6, 2014.

Harrison, S.P., Bartlein, P.J., Izumi, K., Li, G., Annan, J., Hargreaves, J., Braconnot, P., and Kageyama, M.: Evaluation of CMIP5 palaeo-simulations to improve climate projections, Nature Clim. Change, 5, 735-743, https://doi.org/10.1038/nclimate2649, 2015.

Harrison, S.P., Bartlein, P.J., and Prentice, I.C.: What have we learnt from palaeoclimate simulations?, J. Quat. Sci., 31, 363-385, 2016.

Herbert, A.V., and Harrison, S.P.: Evaluation of a modern-analogue methodology for reconstructing Australian palaeoclimate from pollen, Rev. Palaeobot. Palynol., 226, 65-77, https://doi.org/10.1016/j.revpalbo.2015.12.006, 2016.

Hill, D.J., Haywood, A.M., Lunt, D.J., Hunter, S.J., Bragg, F.J., Contoux, C., Stepanek, C., Sohl, L., Rosenbloom, N.A., Chan, W-L., Kamae, Y., Zhang, Z., Abe-Ouchi, A., Chandler, M.A., Jost, A., Lohmann, G., Otto-Bliesner, B.L., Ramstein, G., and Ueda, H.: Evaluating the dominant components of warming in Pliocene climate simulations, Clim. Past, 10, 79-90, DOI:10.5194/cp-10-79-2014, 2014.

Izumi, K., and Bartlein, P.J.: North American paleoclimate reconstructions for the Last Glacial Maximum using an inverse modeling through iterative forward

modeling approach applied to pollen data, Geophys. Res. Letters, 43, doi:10.1002/2016GL070152, 2016.
Izumi, K., Bartlein, P.J., and Harrison, S.P.: Consistent behaviour of the climate system in response to past and future forcing, Geophys. Res. Letters, 40, 1817-1823, DOI:10.1002/grl.50350, 2013.
Izumi, K., Bartlein, P.J., and Harrison, S.P.: Energy-balance mechanisms underlying consistent large-scale temperature responses in warm and cold climates, Clim. Dynam., 44, 3111-3127, DOI: 10.1007/s00382-014-2189-2, 2014.
Jolly, D., and Haxeltine, A.: Effect of low glacial atmospheric CO_2 on tropical African montane vegetation, Science, 276, 786–788, DOI: 10.1126/science.276.5313.786, 1997.
Joussaume, S. and Taylor, K. E.: Status of the Paleoclimate Modeling Intercomparison Project (PMIP), In: Proceedings of the First International AMIP Scientific Conference, WCRP Report, 425–430, 1995.
Jungclaus, J.H., Keenlyside, N., Botzet, M., Haak, H., Luo, J-J., Latif, M., Marotzke, J., Mikolajewicz, U., and Roeckner, E.: Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM, J. Clim., 19, 3952–3972, DOI: 10.1175/JCLI3827.1, 2006.
Kageyama, M., Braconnot, P., Harrison, S.P., Haywood, A.M., Jungclaus, J.H., Otto-Bliesner, B.L., Peterschmitt, J.Y., Abe-Ouchi, A., Albani, S., Bartlein, P.J., Brierley, C., Crucifix, M., Dolan, A., Fernandez-Donado, L., Fischer, H., Hopcroft, P.O., Ivanovic, R.F., Lambert, F., Lunt, D.J., Mahowald, N.M., Peltier, W.R., Phipps, S.J., Roche, D.M., Schmidt, G.A., Tarasov, L., Valdes, P.J., Zhang, Q., and Zhou, T.: The PMIP4 contribution to CMIP6: Part 1: Overview and over-arching analysis plan, Geosci. Model Develop., 11, 1033-1057, https://doi.org/10.5194/gmd-11-1033-2018, 2018.
Kaplan, J.O., Bigelow, N.H., Bartlein, P.J., Christensen, T.R., Cramer, W., Harrison, S.P., Matveyeva, N.V., McGuire, A.D., Murray, D.F., Prentice, I.C., Razzhivin, V.Y., Smith, B. and 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, palaeodata-model comparisons, and future projections, J. Geophys. Res.-Atmosphere 108, No. D19, 8171, DOI: 10.1029/2002JD002559, 2003.
Keenan, T., Serra, J.M., Lloret, F., Ninyerola, M., and Sabate, S.: Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO_2 matters! Glob. Change Biol., 17, 565-579, DOI: 10.1111/j.1365-2486.2010.02254.x, 2011.
Kirtman, B., Power, S.B., Adedoyin, J.A., Boer, G.J., Bojariu, R., Camilloni, I., Doblas-Reyes, F.J., Fiore, A.M., Kimoto, M., Meehl, G.A., Prather, M., Sarr, A., Schär, C., Sutton, R., van Oldenborgh, G.J., Vecchi G., and Wang, H.J.: Near-term climate change: projections and predictability. 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 TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds). Cambridge University Press, Cambridge, UK, 953–1028, 2013.

Lebamba, J., Vincens, A., Maley, J.: Pollen, vegetation change and climate at Lake Barombi Mbo (Cameroon) during the last ca. 33 000 cal yr BP: a numerical approach. Clim. Past, 8, 59-78, https://doi.org/10.5194/cp-8-59-2012, 2012.

Li, G., Harrison, S.P., Bartlein, P.J., Izumi, K., and Prentice, I.C.: Precipitation scaling with temperature in warm and cold climates: an analysis of CMIP5 simulations, Geophys. Res. Letters, 40, 4018-4024, doi:10.1002/grl.50730, 2013.

Li, L., Lin, P., Yu, Y., Wang, B., Zhou, T., Liu, L., Liu, J., Bao, Q., Xu, S., Huang, W., Xia, K., Pu, Y., Dong, L., Shen, S., Liu, Y., Hu, N., Liu, M., Sun, W., Shi, X., Zheng, W., Wu, B., Song, M., Liu, H., Zhang, X., Wu, G., Xue, W., Huang, X., Yang, G., Song, Z., and Qiao, F.: The flexible global ocean-atmosphere-land system model, Grid-point Version 2: FGOALS-g2. Advan. Atmos. Sci., 30, 543-560, https://doi.org/10.1007/s00376-012-2140-6, 2013.

Liu, D. C., Nocedal, J.: On the limited memory BFGS method for large scale optimization. Math. Programming, 45, 503–528, doi: 10.1007/BF01589116, 1989.

Loomis, S. E., Russell, J. M., Verschuren, D., Morrill, C., De Cort, G., Sinninghe Damste, J. S., Olago, D., Eggermont, H., Street-Perrott, F.A., Kelly, M. A.: The tropical lapse rate steepened during the Last Glacial Maximum. Science Advances, 3, e1600815.doi:10.1126/sciadv.1600815, 2017.

Lunt, D.J., Abe-Ouchi, A., Bakker, P., Berger, A., Braconnot, P., Charbit, S., Fischer, N., Herold, N., Jungclaus, J. H., Khon, V. C., Krebs-Kanzow, U., Langebroek, P. M., Lohmann, G., Nisancioglu, K. H., Otto-Bliesner, B. L., Park, W., Pfeiffer, M., Phipps, S. J., Prange, M., Rachmayani, R., Renssen, H., Rosenbloom, N., Schneider, B., Stone, E. J., Takahashi, K., Wei, W., Yin, Q., and Zhang, Z. S.: A multi-model assessment of last interglacial temperatures. Clim. Past, 9, 699–717, DOI: 10.5194/cp-9-699-2013, 2013.

Martin Calvo, M., Prentice, I.C., and Harrison, S.P.: Climate versus carbon dioxide controls on biomass burning: a model analysis of the glacial-interglacial contrast, Biogeosci., 11, 6017–6027, DOI:10.5194/bg-11-6017-2014, 2014.

Martin Calvo, M., and Prentice, I.C.: Effects of fire and CO2 on biogeography and primary production in glacial and modern climates, N. Phytol., 208, 987–994, https://doi.org/10.1111/nph.13485, 2015.

Mauri, A., Davis, B.A.S., Collins, P.M., and Kaplan, J.O.: The influence of atmospheric circulation on the mid-Holocene climate of Europe: a data-model comparison. Climate of the Past 10: 1925–1938, DOI: 10.5194/cp-10-1925-2014, 2014.

Menke, W.: Geophysical data analysis: Discrete inverse theory (Matlab 3rd ed.). Cambridge, Massachusetts: Academic Press, 2012.

Deleted: ¶

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

Nichols, N.K.: Mathematical concepts of data assimilation, In: Lahoz, W., Khattatov, B., and Menard, R. (Eds.), Data Assimilation, Springer, 2010.

Perez-Sanz, A., Li, G., Gonzalez, P., and Harrison, S.P.: Evaluation of seasonal climates of northern Africa and the Mediterranean in the CMIP5 simulations, Clim. Past, 10, 551-568, DOI:10.5194/cp-10-551-2014, 2014.

Prentice, I.C., Dong, N., Gleason, S.M., Maire, V., and Wright, I.J.: Balancing the costs of carbon gain and water loss: testing a new quantitative framework for plant functional ecology, Ecol. Letters, 17, 82-91, https://doi.org/10.1111/ele.12211, 2014.

Prentice, I.C., Cleator, S.F., Huang, Y.H., Harrison, S.P., and Roulstone, I.: Reconstructing ice-age palaeoclimates: Quantifying low-CO_2 effects on plants, Glob. Planet. Change, 149, 166-176, https://doi.org/10.1016/j.gloplacha.2016.12.012, 2017.

Schmidt, G.A., Ruedy, R., Hansen, J.E., Aleinov, I., Bell, N., Bauer, M., Bauer, S., Cairns, B., Canuto, V., Cheng, Y., Del Genio, A., Faluvegi, G., Friend, A.D., Hall, T.M., Hu, Y., Kelley, M., Kiang, N.Y., Koch, D., Lacis, A.A., Lerner, J., Lo, K.K., Miller, R.L., Nazarenko, L., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Russell, G.L., Sato, M., Shindell, D.T., Stone, P.H., Sun, S., Tausnev, N., Thresher, D., and Yao, M-S.: Present day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite and reanalysis data, J. Clim., 19, 153-192, https://doi.org/10.1175/JCLI3612.1, 2006.

Roderick, M.L., Greve, P., and Farquhar, G.D.: On the assessment of aridity with changes in atmospheric CO_2, Water Resour. Res., 51, 5450–63, DOI:10.1002/2015WR017031, 2015.

Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M-P., Planton, M.S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F.: The CNRM-CM5.1 global climate model: description and basic evaluation, Clim Dyn., 759, doi:10.1007/s00382-011-1259-y, 2012.

Wang, H., Prentice, I.C., and Davis, T.W.: Biophysical constraints on gross primary production by the terrestrial biosphere, Biogeosci., 11, 5987-6001, https://doi.org/10.5194/bg-11-5987-2014, 2014.

Wang, H., Prentice, I.C., Cornwell, W.M, Keenan, T.F., Davis, T.W., Wright, I.J., Evans, B.J., and Peng, C.: Towards a universal model for carbon dioxide uptake by plants, Nature Plants, 3, 734–741, https://doi.org/10.1038/s41477-017-0006-8, 2017.

Wang, Y., Herzschuh, U., Shumilovskikh, L. S., Mischke, S., Birks, H. J. B., Wischnewski, J., Bohner, J., Schlutz, F., Lehmkuhl, F., Diekmann, B., Wunnemann, B., Zhang, C.: Quantitative reconstruction of precipitation changes on the NE Tibetan Plateau since the Last Glacial Maximum – extending the concept of pollen source area to pollen-based climate reconstructions from large lakes, Clim. Past, 10, 21-39, https://doi.org/10.5194/cp-10- 21-2014, 2014.

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori., S., Kawamiya, M.: MIROC-ESM: model description and basic results of CMIP5-20c3m experiments, Geosci. Mod. Dev., 4, 845-872, https://doi.org/10.5194/gmd-4-845-2011, 2011.

Wu, H., Guiot, J., Brewer, S., 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, 211–229, DOI: 10.1007/s00382-007-0231-3, 2007.

Wu, H., T. Guiot, J., Peng, C., and Guo, Z.: New coupled model used inversely for reconstructing past terrestrial carbon storage from pollen data: Validation of model using modern data, Glob. Change Biol., 15, 82–96, https://doi.org/10.1111/j.1365-2486.2008.01712.x, 2009.

Yukimoto, S., Yoshimura, H., Hosaka, M., Sakami, T., Tsujino, H., Hirabara, M., Tanaka, T.Y., Deushi, M., Obata, A., Nakano, H., Adachi, Y., Shindo, E., Yabu, S., Ose, T., Kitoh, A.: Meteorological Research Institute-Earth System Model v1 (MRI-ESM1) – Model Description, Tech. Rep. Meteor. Res, Inst., 64, 88 pp., http://www.mri-jma.go.jp/Publish/**Technical**/DATA/**VOL_64**/index.html, 2011.

Zhang, L., Hickel, K., Dawes, W.R., Chiew, F.H.S., Western, A.W., Briggs, P.R.: A rational function approach for estimating mean annual evapotranspiration. Water Resources Research 40, W02502, DOI:10.1029/2003WR002710, 2004.

Figures and Tables Captions

Figure 1: The distribution of the site-based reconstructions of climatic variables at the Last Glacial Maximum. The individual plots show sites providing reconstructions of (a) moisture index (MI), (b) mean annual precipitation (MAP), (c) mean annual temperature (MAT), (d) mean temperature of the coldest month (MTCO), (e) mean temperature of the warmest month (MTWA) and growing degree days above a baseline of 5◦ C (GDD5). The original reconstructions are from Bartlein et al. (2011) and Prentice et al. (2017).

Figure 2: Uncertainties associated with the climate prior. The climate is derived from a multi-model mean of the ensemble of models from the Palaeoclimate Modelling Intercomparison Project (PMIP) and is shown in SI Figure 1. The uncertainties shown here are the standard deviation of the multi-model ensemble values. The individual plots show the variance for the simulated (a) moisture index (MI), (b) mean annual precipitation (MAP), (c) mean annual temperature (MAT), (d) mean temperature of the coldest month (MTCO), (e) mean temperature of the warmest month (MTWA) and growing degree days above a baseline of 5◦ C (GDD5).

Deleted: uncertainties

Figure 3: Uncertainties on the analytical reconstructions. These uncertainties represent a combination of the uncertainty on the site-based reconstructions, and the grid-based variance on the prior and the global variance from the prior.

Deleted: These uncertainties represent a combination of the errors on the site-based reconstructions, and the grid-based errors on the prior and the global uncertainty from the prior.

Figure 4: Analytically reconstructed climate, where areas for which the site-based data provide no constraint on the prior have been masked out. The individual plots show reconstructed (a) moisture index (MI), (b) mean annual precipitation (MAP), (c) mean annual temperature (MAT), (d) mean temperature of the coldest month (MTCO), (e) mean temperature of the warmest month (MTWA) and growing degree days above a baseline of 5◦ C (GDD5). The anomalies are expressed relative to the long term average (1960-1990) values from the Climate Research Unit (CRU) historical climatology data set (CRU CL v2.0 dataset, New et al., 2002).

Figure 5: Impact of CO_2 on reconstructions of moisture-related variables. The individual plots show (a) the change in moisture index (MI) and (b) the change in mean annual precipitation (MAP) compared to the original pollen-based reconstructions for the LGM before (circles) and after (crosses) the physiological impacts of [CO_2] on water-use efficiency are taken into account. The third plot (c) shows the relative difference in MI and MAP as a result of [CO_2], shown as the percentage difference between the no-[CO_2] and [CO_2] calculations.

Deleted: Impact of CO_2 on reconstructions of moisture-related variables. The individual plots show (a) the change in moisture index (MI) and (b) the change in mean annual precipitation (MAP) when the physiological impacts of [CO_2] on water-use efficiency are taken into account. The third plot (c) shows the relative difference in MI and MAP as a result of [CO_2], shown as the percentage difference between the no-[CO_2] and [CO_2] calculations.

Table 1: Details of the models from the Palaeoclimate Modelling Intercomparison Project (PMIP) that were used for the Last Glacial Maximum (LGM) simulations used to create the prior.

Table 1: *Details of the models from the third phase of the Palaeoclimate Modelling Intercomparison Project (PMIP3) that were used for the Last Glacial Maximum (LGM) simulations used to create the prior. Coupled ocean-atmosphere models are indicated as OA, which OAC models have a fully interactive carbon cycle. The resolution in the atmospheric, oceanic and sea ice components of the models is given in terms of numbers of grid cells in latitude and longitude.*

Model name	Type	Resolution			Year length	Reference
		Atmosphere	Ocean	Sea Ice		
CCSM4	OA	192, 288	320, 384	320, 384	365	Gent et al. (2011)
CNRM-CM5	OA	128, 256	292, 362	292, 362	365-366	Voldoire et al. (2012)
MPI-ESM-P	OA	96, 192	220, 256	220, 256	365-366	Jungclaus et al. (2006)
MRI-CGCM3	OA	160, 320	360, 368	360, 368	365	Yukimoto et al. (2011)
FGOALS-g2	OA	64, 128	64, 128	64, 128	365	Li et al. (2013)
COSMOS-ASO	OAC	96, 48	120, 101	120, 101	360	Budich et al. (2010)
IPSL-CM5A-LR	OAC	96, 96	149, 182	149, 182	365	Dufresne et al., 2013
MIROC-ESM	OAC	64, 128	192, 256	192, 256	365	Watanabe et al. (2011)

Appendix

We define e as the water lost by transpiration (E) per unit carbon gained by photosynthesis (A). This term, the inverse of the water use efficiency, is given by:

$$e = E/A = 1.6\,D\,/\,((1 - \chi)\,c_a) \qquad \text{(A1)}$$

where D is the leaf-to-air vapour pressure deficit (Pa), c_a is the ambient CO_2 partial pressure (Pa) and χ is the ratio of leaf-internal CO_2 partial pressure (c_i) to c_a. An optimality-based model (Prentice *et al.* 2014), which accurately reproduces global patterns of χ and its environmental dependencies inferred from leaf $\delta^{13}C$ measurements (Wang *et al.* 2017), predicts that:

$$\chi = (\Gamma^*/c_a) + (1 - \Gamma^*/c_a)\,\xi/(\xi + \sqrt{D}) \qquad \text{(A2a)}$$

and

$$\xi = \sqrt{(\beta(K + \Gamma^*)/1.6\,\eta^*)} \qquad \text{(A2b)}$$

where Γ^* is the photorespiratory compensation point of C_3 photosynthesis (Pa), β is a constant (estimated as 240 by Wang *et al.* 2017), K is the effective Michaelis-Menten coefficient of Rubisco (Pa), and η^* is the ratio of the viscosity of water (Pa s) at ambient temperature to its value at 25˚C. Here K depends on the Michaelis-Menten coefficients of Rubisco for carboxylation (K_C) and oxygenation (K_O), and on the partial pressure of oxygen O (Farquhar *et al.* 1980):

$$K = K_C\,(1 + O/K_O) \qquad \text{(A3)}$$

Standard values and temperature dependencies of K_C, K_O, Γ^* and η^* are assigned as in Wang *et al.* (2017).

The moisture index MI is expressed as

$$\mathrm{MI} = P/E_q,\; E_q = \sum\nolimits_n (R_n/\lambda)\,s/(s + \gamma) \qquad \text{(A4)}$$

where P is annual precipitation, R_n is net radiation for month n, λ is the latent heat of vaporization of water, s is the derivative of the saturated vapour pressure of water with respect to temperature (obtained from a standard empirical formula also used by Wang *et al.* 2017), and γ is the psychrometer constant. We assume that values of MI reconstructed from fossil pollen assemblages, using contemporary pollen and climate data either in a statistical calibration method or in a modern-analogue search, need to be corrected in such a way as to preserve the contemporary relationship between MI and e, while taking into account the change in e that is caused by varying c_a and temperature away from contemporary values. The sequence of calculations is as follows. (1) Estimate e and its derivative with respect to temperature ($\partial e/\partial T$) for the contemporary c_a and climate, using equations (A1) – (A3) above. (2) Use the e and $\partial e/\partial T$ *to* calculate $\partial D/\partial T$ given the palaeo c_a (measured in ice-core data) and temperature (reconstructed from pollen data), via a series of analytical equations that relate $\partial e/\partial T$ to $\partial D/\partial T$ and hence to s. (3) Use the new $\partial D/\partial T$ *and relative humidity (from the PMIP3 average)* to derive a new value of s. (4) Re-calculate MI using a palaeo estimate of R_n (modelled as in Davis et al., 2017) and the new value of s.

Formatted: Font: Italic