



**UCD School of  
Mathematics and Statistics**

Room G09, Science Centre – North  
University College Dublin  
Belfield, Dublin 4, Ireland

Tel +353 1 716 2422  
Fax +353 1 716 1172

**Scoil na  
nEolaíochtaí Matamaitice UCD**

Seomra G09, Lárionad Eolaíochta Uí Bhriain - Thuaidh  
An Coláiste Ollscoile, Baile Átha Cliath  
Belfield, Baile Átha Cliath 4, Éire

Email [Andrew.Parnell@ucd.ie](mailto:Andrew.Parnell@ucd.ie)  
Web [mathsci.ucd.ie/people/parnell\\_a](http://mathsci.ucd.ie/people/parnell_a)

September 5, 2016.

Dear Professor Carrión,

Many thanks for the detailed comments regarding our paper titled ‘Joint palaeoclimate reconstruction from pollen data via forward models and climate histories’. We have addressed all of the queries for the two referees for whom you forwarded responses. Just as we were preparing our final submission we received a further review via a personal communication from Michel Crucifix. We have incorporated all of his very useful comments except in one case where he suggested a substantial revision of one of the figures. We feel that these extra comments further improve the quality and reach of the paper.

We hope that you now find the paper suitable for publication. If you have any further queries please do not hesitate to get in contact with me.

Yours sincerely,

Prof Andrew Parnell CStat

## Comments from Reviewer number 1:

This important paper presents to the Quaternary community a recent development, where three independent models are combined to provide a fascinating new approach to go from pollen counts to climate reconstruction. Uncertainties are calculated by combining those within the three modules and by assessing changes between time-steps instead of just those of individual depths - i.e. looking at actual dynamics instead of at single slices of time. The method is applied on two published case-studies, showing interesting new insights. Pending some clarifications I would suggest accepting this manuscript for publication.

The ms describes most of the framework in clear words, leaving the detail in more technical papers already published in statistics journals (Salter-Townshend and Haslett 2012, Parnell et al. 2015). Some strong but correct remarks are made regarding the failures of existing methods commonly applied by the Quaternary research community. The ms convincingly argues that this community needs to shift toward using forward, Bayesian models.

We are very grateful to this reviewer for their comments.

My main fundamental criticism is that currently not enough detail is provided to properly follow the first component of Bclim. How is modern pollen used to interpret the fossil pollen in terms of climate, and which non-climate processes could be messing things up? The reader is referred to Parnell et al. 2015 and Salter-Townshend & Haslett 2012, but neither of these papers describe the required and fundamental mechanisms in sufficient detail for Quaternary-oriented readers.

1. Forestation (e.g. plantations of non-native trees), deforestation and other major vegetation/soil changes over the past centuries will cause recent pollen spectra to be highly disturbed by a range of important, additional, non-climate factors; how is this taken into account?
2. Not all pollen fluctuations are (directly) forced by climate change, and to assume so would be a crude over-simplification. Yes, there could be lags, but other more fundamental factors are at work. For example, over longer time-scales such as that over which the sites accumulated, large-scale migration of plant species has been important.

As with all palaeoenvironmental reconstructions based upon variations in relative abundance of taxa that result from their differing environmental niches, our reconstructions depend upon surface samples. The referee's first concern relates to the extent to which pollen surface samples

may be biased by non-climatic influences, rather than giving a signal related primarily to the response of the pollen-producing plants to climate. Clearly such biases will be present, although of those suggested by the referee, deforestation, leading to a reduction in the proportion of tree pollen in surface samples, is the only one of regional significance. Planting of non-native trees is of limited importance in most regions of the northern hemisphere from which our surface samples are derived, and it is arguable that across most of these areas “other major vegetation/soil changes over past centuries” are mostly responses to climatic changes during that period. In any case, pollen spectra from higher latitudes are mostly dominated by tree pollen, and whilst human activities have reduced overall forest cover in some regions (e.g. western Europe), they have had a much lesser impact upon the composition of the remaining forests. As a result, a clear climatic signal is apparent in the varying composition of pollen surface samples, and is the dominant signal at regional scales. In order to ensure that it is this regional signal that we calibrate, we exclude pollen surface samples taken from beneath forest canopies or from very small basins that are dominated by local signals that more often may be non-climatic.

The referee’s second concern appears to focus upon the long-term and large-scale migrations of plant species, these implicitly being argued by the referee not to be responses to climatic change. This argument conflicts with a great deal of evidence, published especially over the last three decades, that provides a compelling case that the large-scale migrations of plants, and of other groups of organisms, over glacial-interglacial time scales during the Pleistocene, were at least predominantly responses to climatic changes. For example, Prentice et al. (1991; *Vegetation and climate change in Eastern North America since the last glacial maximum. Ecology*, 72, 2038-56.) provide clear evidence that the migrations of various eastern North American trees since the last glacial maximum were responses to climatic changes. Thus, notwithstanding arguments by some to the contrary, the distributions of species, and thus vegetation composition and pollen spectra, are in dynamic equilibrium with the changing climate.

In response to these important comments, we have added the following paragraph at the end of Section 3.1 to clarify:

In this version of Bclim, as in all modelling, the approximations are simplifications of reality. In particular, vegetation composition may change in response to changes in other climatic or non-climatic factors, potentially leading to biases in the modelled relationships and hence in the reconstructions. In some regions, and during more recent time, anthropogenic interference with the vegetation may have led to changes which the model will falsely attribute to climatic change. More generally, our models do not account for potential lags in the responses of vegetation to climatic changes that may result from limitations upon species’ migration rates and/or the requirement for a disturbance episode to facilitate both these migrations and the associated changes of vegetation composition (see e.g. Bradshaw & Zackrisson, 1990). The model assumes that changes in vegetation, and hence in the pollen recruited to sediments, occur immediately following a change in climate. Given the evidence that the lags in vegetation response are generally much less than temporal uncertainties associated with the use of  $^{14}\text{C}$  dating (see e.g. Tinner & Lotter, 2001; Williams et al., 2002), we believe that this is an acceptable simplification. Furthermore, we envisage that richer approximations, accounting both for the influences of other factors and for the mechanisms of ecological processes, will be a major focus of future

research.

3. Other queries:

- (a) Which pollen types are most important in your modelling process?
- (b) How do you deal with the dependence of pollen counts on other counts (percentage dependance)?
- (c) Pollen types that are included in your model could still be influenced by fluctuations in non-included pollen.
- (d) Are the pollen counts percentage data, or concentration?

- (a) The most important pollen types will depend on the location in climate space (the region defined by GDD5/MTCO/AETPET). If, for example, a particular taxon is only found in a very specific part of climate space then, even if it is not very abundant, it will have a very strong say in the final climate estimates.
- (b) The percentage dependence between pollen counts is handled via the Multinomial framework of Salter-Townshend and Haslett (2012), and Sweeney (2012). We have purposefully not included too many technical details on this work as we feel it might detract readers from the broader messages in the paper.
- (c) The 28 pollen taxa selected for inclusion in our model include most of those that regularly comprise substantial proportions of the pollen rain in the higher latitudes of the northern hemisphere, and all of those that do so in northern Eurasia. We also include a number of taxa that make smaller contributions, but that are strongly indicative of particular climatic conditions. As a result, the 28 taxa together account for the majority of pollen in most surface and fossil pollen spectra used. Their proportions will thus only very rarely be substantially influenced by other pollen taxa that we do not include.
- (d) Pollen data can be provided as counts or percentages. Ideally counts should be used as these contain the most information but the ability to handle different count sizes for each layer is not implemented in the current version of Bclim.

To help the reader with these important issues, we have added the following paragraph in the Discussion:

It is important to separate out the ideas presented in this paper (forward modelling, joint inference, climate histories) from the practicalities of the software Bclim. The former is broad, extendible, and applicable to a wide variety of reconstruction techniques. The latter at present is still simplistic in its nature. For example whilst the 28 pollen taxa used accounts for the majority of pollen in most surface and fossil pollen spectra used, larger quantities or different proxies could be combined into a forward model. Similarly the modern data and statistical models used by Bclim for calibrating pollen samples could better model the proportion dependence induced by the compositional nature of pollen counting. All these are avenues for future research.

4. Is the same modern pollen dataset used for both case-studies, and if so, is this justifiable given that they come from two quite different regions?

The same modern pollen dataset is used for both case studies, and would be used for every fossil pollen dataset presented to Bclim. This can be justified on several grounds, but in particular:

1. the dataset comprises 7815 surface samples from throughout the higher latitudes of the northern hemisphere; and
2. the palaeovegetation at a site will often differ from any found today in the surrounding region and is thus more likely to be ‘matched’ correctly using such an extensive surface sample compilation rather than one restricted to the region surrounding the site.

As currently configured, Bclim is thus appropriate for application to fossil pollen datasets from the higher latitudes of the northern hemisphere. Given further appropriate datasets of modern samples, however, future work could extend this scope both to other regions and/or to other biological proxies.

5. Could diagrams of the main pollen records for both sites be provided as supplementary information?

These have now been included in Appendix 2 with references to the appendix in the appropriate parts of Section 5.

Minor comments:

- (a) Abstract: at first mention of climate histories please call them dynamic climate histories - which will call further attention to the move away from snapshots
- (b) Fig. 1, top panel. There is a white blob around 16 kcal BP (c. 4500-6500 GDD5). Is this real or rather indicating a problem with drawing the contours?
- (c) Why do the climate ribbons trumpet out so much towards the extremes (especially 0 kcal BP)? Could the use of assumed core surface ages (e.g., 0cm=AD 1995 +/- 10) help constraining the reconstructions?
- (d) Some of the climate ribbon at the bottom panel of Fig. 1 extends beyond 1.0 AET/PET but none of the blue blobs do. Is this realistic?

Thanks for spotting all of these. In response:

- (a) Done
- (b) Yes there seemed to be a small issue in drawing the contours. We have re-written the code slightly to fix this. However, outlying blobs are a feature of the way the climate clouds are plotted (as highest density regions) since they are multimodal and based on a finite set of samples, rather than a full probability distribution.
- (c) The climate ribbons trumpet out due to the random walk nature of the stochastic volatility model (SVM) we use to constrain climate over time. Future versions of Bclim will allow both the surface ages and the climates to be constrained.
- (d) The SVM we use is based on an assumed normally distributed climate increment between time steps. Whilst this works well for MTCO (which is unbounded), it is slightly unsatisfactory for GDD5 (bounded at 0) and especially for AET/PET (bounded above and below at 0 and 1). We could, for example, run the SVM on the logit of the AET/PET but this would then disallow the values 0 and 1 which are perfectly feasible values of the climate variable. A further alternative is to project values above 1 or below 0 to be exactly 1 or 0 respectively, though this is somewhat ad-hoc. We hope to find a more satisfactory solution in the near future

To address the first issue we have updated Figures 1 and 3. To address the latter two issues we have added the following paragraph in Section 5.1:

There are some features of Figure 1 which are clearly unrealistic, and are side-effects of the first generation, proof-of concept nature of our Bclim model. The first of these is the trumpeting out of uncertainty as we reach the present. Future versions of Bclim could allow the user to input constrained climate values for the present. Another is the possibility for the climate ribbon to lie outside the plausible range of the climate variable (see bottom panel of Figure 6 period 0 to 1k BP). This arises from the normally-distributed nature of the stochastic volatility climate model discussed in Section 4.3. Again, future versions may allow for a richer climate model surrogate that respects the bounds of the climate variable in question. We ignore these issues in our remaining interpretations.

### Comments from Reviewer number 3:

This paper provides a fully formed framework (with software) for proxy-based joint reconstruction of multivariate past climate; it is illustrated using pollen data from two lake sediment cores and three dimensions of climate. Jointly reconstructing climate in this way, using forward models to represent the processes that gave rise to the proxy records, allows users to draw strength from all of the data they are analysing and thus, in many cases, to reduce the resulting uncertainty on the climate reconstructions. This increase in formality and resulting decrease in uncertainty is likely to prove popular with the user community and so I foresee this paper being influential and highly cited. I am convinced about this, not just because I think the science is well conceived and implemented, but also because the paper is beautifully crated. Indeed, I would go so far as to describe both the work itself and this paper as a tour de force. The paper is highly tailored to the intended readership and densely packed with key concepts, explanations and illustrations. I found no part of the prose or supporting material that I would suggest removing or substantially revising. I have just a few minor comments that I hope might help the authors as they finalise the manuscript for publication.

We really appreciate these comments and hope that the readers of QSR feel the same.

I wonder whether the term “forward model” needs explaining earlier in the paper. My feeling is that most readers who have encountered this term in the past will have done so in the context of mechanistic forward models which seek to represent the means by which actions in one part of a system cause changes in another. Such models are usually deterministic and rely on systems of differential equations which can be hard to invert. I am concerned that some readers will hold this image of a forward model fairly firmly in their minds and thus may struggle to grasp the relative simplicity of what’s offered here. I think it would help readers to be told earlier on than at present that, to be useful in the present context, forward models need not be fully mechanistic and that (as illustrated) simple descriptive forward models can be very powerful.

We have added an extra paragraph in the introduction to make these points clear:

We see a forward model (sometimes known as a proxy systems model, Evans et al., 2013) as the causal chain from which climate is transformed into proxy data stored in an archive. Our definition is broad, encompassing both deterministic and statistical approaches (ideally combined together), but with a clear focus on accounting for uncertainty at each stage. This uncertainty may be due to unknown

processes which lend themselves to deterministic modelling, such as the means by which pollen is spread through the local area (e.g. Garreta et al., 2009), to stochastic processes such as the probability of detecting a particular variety of pollen through a microscope, given that it is present. In our approach we combine the forward model with a simple stochastic climate model via Bayesian inference, which allows us to produce climate histories with narrower uncertainties than using climate dimensions and pollen slices individually.

The term “archive” is not defined. It is used at the top of page 6 for the first time, but it’s a useful concept that merits introduction earlier on.

We have now introduced the term in the introduction as above, and have expanded the text in Section 3.1 to remind readers of its use.

There is potential for confusion in the authors’ use of the term “layer”. In some places it is clearly referring to a physical layer in a laminated sediment, in other places to a slice from a core and elsewhere there is potential for confusion since it could refer to part of a hierarchical model. I suggest that it be reserved only for the first of these and that other terms be used elsewhere. For sections of a core I recommend that “slice” be adopted as the generic term since this works even if the slice constitutes one layer from a laminated sediment.

We have changed all occurrences of ‘layer’ to ‘slice’ in the text, except where they refer to laminated sediments. We have further changed all occurrences in the Bclim package, and updated all plots accordingly

The term “layer-by-layer” is used which may need to become slice-by-slice given recommendation 3). I also recommend hyphenating it as indicated. At present hyphenation is inconsistent.

Done

To a statistician, it is obvious why joint modelling leads to reduced uncertainties in the reconstructions, but I think it might be worth spelling out for the QSR reader why this is. Phrases like “These intervals will be narrower than those generated by ...” (page 22,) without telling us (at least intuitively) why, are



worth reconsidering. I would explain this early on and then refer back to this benefit at points like page 22. This seems particularly important since it is a highlight of the summarised benefits in the abstract.

We have tried to make this point more clearly. It is still made in the Abstract and in Section 6.1, but is now also mentioned in the Introduction in the new paragraph created above, in Section 2 where we discuss slice-by-slice inference, and finally in Section 4.3 where we discuss the creation of climate histories.

I am a bit surprised that the paper doesn't have an acknowledgements section. I have been present at several meetings and workshops that I feel sure must have influenced the authors' thinking on these topics in particular the SUPRAnet meetings (funded by Leverhulme) and the Isaac Newton Institute workshop on Uncertainty in Climate Prediction: Models, Methods and Decision Support.

This was an oversight on our part. An acknowledgements section containing the above has now been included.

Specific suggestions:

1) Final bullet point at the top of Page 2: "may or may not form a subset of the layers". Strictly the dates don't form the layers so I suggest "may or may not be from a subset of the layers".

Changed

2) End of second paragraph Page 2: "play with" trivialises what's offered here. I suggest "explore".

Changed

3) Last paragraph before Section 2 (Page 4): I had trouble parsing the end of the sentence that starts "The current paradigm ...". I think the structure needs looking at, but at the same time I suggest that the authors consider mentioning

comparing as well as combining of reconstructions since even the former is not achievable without coherent error estimation.

We have changed this sentence to:

At a more subtle level, the current paradigm, where uncertainty is encapsulated in published error bars (or, worse, in a single RMSEP), provides no basis for comparing, let alone combining, reconstructions which may be similarly uncertain.

4) First paragraph of Section 3.1 (Page 5): I found the choice of reference for the use of forward models in tree-ring dating a little odd. Towlinski-Ward et al (2011) seems an arbitrary choice from a substantial body of literature. I suggest either explaining the choice (eg that TW11 offers a simple, invertible model) or adding in one or two of the earlier references on which TW11 builds.

We have clarified that we are using this reference because of its direct applicability to our proposed approach.

5) Table 2 is potentially misleading or confusing since it suggests a one-step modelling process when, immediately below, we are told that the approach is a sequential, step-based one. Surely Table 2 would be more useful if it showed the steps eg as a series of connected boxes with appropriate labels.

We have removed Table 2 to avoid this confusion.

6) First paragraph Section 3.3 (Page 7): is the most serious example of potential confusion over the term “layer”. Here the authors are discussing a stepwise approach to modelling and say that they are “... including some previously hidden layers which make up an intermediate step in our model”. The next sentence makes it explicit that the layers being referred to are slices (as defined in my suggested terminology above), but the image of a layer within the stepwise model has already been created in the readers mind. I recommend rewording along the following lines to avoid any possible confusion: “... by including the locations of the slices from which the pollen data were obtained. The locations of the slices are an explicit part of the model, but were previously hidden when we considered just the resulting gridded output.”. For this to work, of course, the term “slice” should also replace “layer” in the rest of the paper as suggested above.

This has been changed as suggested.

7) End of first paragraph Section 3.3 (Page 7): the authors' use of the term "data set". Here the authors are summarising model output (in the form of samples) which they are comparing to data. I would prefer to see a comparison made between their model output and "any large sample" since samples of data and samples of model output can usefully be treated in similar ways. I am concerned that the current phrasing might reinforce the sloppy use of language in this community whereby model output are referred to as data and then subjected to inappropriate further processing and modelling.

Since this sentence was simply re-stating points already made in the preceding paragraph, we have removed the offending text.

8) First paragraph Section 4 (Page 9): "taken" should be "take".

Changed

9) Page 9: second sentence of point (3) in numbered list. Suggested rewording for clarity and completeness: "This constrains the climate cloud to exhibit only 'reasonable' climate change given the chronological histories and the prior distribution on climate."

Changed

10) Paragraph immediately before Section 4.1 (Page 9): why 28 taxa? Can these be chosen by the user? Can more be added or deleted?

Sadly not at present. These 28 taxa are fixed and hard-coded into the programme. Future versions may make this more flexible. Since this was also queried by Referee 1, we have added the following paragraph in the discussion:

It is important to separate out the ideas presented in this paper (forward modelling, joint inference, climate histories) from the practicalities of the software Bclim. The

former is broad, extendible, and applicable to a wide variety of reconstruction techniques. The latter at present is still simplistic in its nature. For example whilst the 28 pollen taxa used accounts for the majority of pollen in most surface and fossil pollen spectra used, larger quantities or different proxies could be combined into a forward model. Similarly the modern data and statistical models used by Bclim for calibrating pollen samples could better model the proportion dependence induced by the compositional nature of pollen counting. All these are avenues for future research.

11) Paragraph immediately before Section 4.1 (Page 9): “dating evidence” not “data”. At the end of the sentence on the chronological inputs the authors say “... each of which represents a sample chronology that is consistent with the data”. I would prefer to see this sentence end with “dating evidence” since this is more explicit and makes it clear which part of the data is relevant here.

Changed

12) Page 10: use of term “blob”. I think that this is a useful term, but initially could not see the merit of both “cloud” and “blob”. I wonder whether a slight rewording might help eg “... as a set of blobs, one for each year; given the uncertainties, these overlap to form clouds”.

We have removed the term ‘blob’ to avoid duplicating jargon.

13) First paragraph Page 12: last sentence needs citation to paper where more detail on technicalities of prior specification are offered.

We have added a reference here.

14) Second sentence of last paragraph of Section 4 (Page 12): suggested rewording. “We also plot the individual layer climate clouds as these provide some information about the evidence for specific climate values provided by each layer in the core.”.

Changed

15) Second sentence Section 5.1 (Page 12): splice comma should be a semi-colon.

Changed

16) Third sentence Section 5.1 (Page 12): “sediment samples” should be “sediment slices” (or sediment layers) to match terminology in modelling sections of the paper.

Changed

17) First sentence of first full paragraph on Page 13: why are Bclim and Bchron referred to in this order? You can’t run BClim without first running BChron so I would have them the other way around.

Changed

17) Penultimate sentence of first full paragraph on Page 13: use of term “beyond” implies outside. Suggest rewording to “... climate histories actually look and where they lie within the climate ribbon summary”.

Changed

18) Page 13: concept of signal to noise ratio. I’m not sure how familiar readers of QSR will be with this term. Might it be worth spelling out how the SNR is derived?

We have added in a sentence to give the definition of the SNR

19) Last paragraph on Page 13: “AET/PET furthers shows...” should be “AET/PET further shows...”.

Changed

20) Last paragraph of Section 5.1 (Page 15): “clear” and “clearly” in the same sentence, might be worth rewording.

The first ‘clear’ has been removed

21) First paragraph of Section 6.1: “The ribbon plots will the ...” should be “The ribbon plots will be the ...”.

Changed

22) Penultimate paragraph on Page 22: “...Figure 4 once again raise insights...” should be “...Figure 4 once again raises insights...”.

Changed

23) Second sentence of last paragraph on Page 22: I recommend parenthetic commas around “for example”.

Changed

24) Penultimate sentence of paper (Page 23): ends “anticipated” as does the preceding one. Worth rewording to avoid repetition.

The first ‘anticipated’ sentence has been removed as redundant.

**Comments from late review from Michel Crucifix:**

This article is a companion to Parnell et al., 2015, published in the Journal of the Royal Statistical Society, and to the R package ‘BClim’, available on the R-user community platform CRAN (and also available on the version controlling platform github). The specific objective of this article seems to be to bring the concepts behind BClim to the audience of Quaternary scientists (with a focus on palynology), a natural objective since they constitute the user target of BClim. Mathematical notations and equations are avoided (there is no equation except for the conceptual representation of Bayes’ theorem, p. 6). The statistical concepts are illustrated with a ‘toy example’ and, then, BClim is applied to two real-world, relevant examples. These case studies are analysed with quite some care, enough at least to make them interesting in their own right.

The text is overall easy to read and the concepts are clearly explained. There is quite some emphasis on the fact that the approach is “entirely new” (pp. 4 & 21), and that the authors offer here in a “new paradigm” (p. 1). So it seems that this review should examine the relevance of that claim for novelty, whether the paper makes a good job at bringing statistical concepts to a broader audience, and discuss the scientific accuracy and consistency of the case studies. There will be some critique, but overall the article will be a good contribution to Quaternary Science Reviews and I recommend its publication.

We thank Michel for his very thorough review.

BClim is a Bayesian inversion algorithm, using pollen data counts as inputs to sample a posterior of climate histories, given a likelihood encoding constraints about the climate dynamics and the relationship between climate and pollen counts. BClim may also be (and is generally meant to be) associated with the BChron algorithm to constrain the relationship between depth and time.

In fact quite a bit of literature has flourished in the recent years tackling palaeoclimate inversion problems, accounting for the data generation process and time constraints. The nature and interest of various contributions in this domain will depend on the nature of the observations, the idiosyncrasies and level of sophistication of the data-generating process. The article of Holmstrom (2015) cited by these authors is systematic enough in listing these recent contributions. They include among others the BARCAST model (Werner and Tingley, 2015, Climate of the Past, for one of the most recent ones), which is uncomfortably left un-cited in the present contribution.

We have now included a citation to the BARCAST model.

In the statistical literature, state estimation, given knowledge about a time process is commonly known as smoothing, a task which may be more or less involved (and involving more or less approximations) depending on the dimensionality of the problem and the complexity of the time process. There are even some recent attempts in the palaeoclimate literature to consider stochastic non-linear dynamical systems. Others have simply considered a Gaussian processes (of which the random walk may be seen as a particular case). So time smoothing as such is not new as such. A specificity of the present contribution is to link local climate volatility with that reconstructed in Greenland.

Climate state estimation methods are often quite tailored for specific case studies, and because of their technicality they remain under-appreciated by the Quaternary Science Community. It may be that the revolution the authors seem to be after here, is more about the way of communicating uncertainties, than about a ‘novel’ methodology to the problem of climate reconstruction.

Michel is correct in stating that some of the ingredients here have appeared in simpler forms in other papers. In fact, in each case where we have used the word ‘novel’ or ‘paradigm’ it is with respect to the use and distribution of climate histories, and mentioned as such in the same sentence. We stand by these statements. The first reference to ‘entirely new’ is qualified in a similar way. However, the second reference to ‘entirely new’ in the Discussion has been reduced in emphasis.

Regarding the methodology, precisely, the authors refer the reader to the Journal of the Royal Statistical Society paper. The naive reader might understand that the inversion problem is exclusively solved by brute-force Monte-Carlo, while precisely the statistician’s art goes into adopting clever approximations depending on the scientific needs and the nature and abundance of the data. This is the reason why there are different algorithms for different models. This message ends up a bit buried in the present article.

We have tried not to put off the reader with complicated statistical concepts. We hope that those who really wish to get into the bowels of the models will read the RSS paper, and so we have referenced this repeatedly.

Finally, near the end of the article, the authors evoke (indirectly) GCMs as they cite PMIP, arguably a bit as if BClim could straightforwardly be extended to accommodate such models. We all know quite well that quite a number of considerations need to be addressed to sample GCM palaeoclimate histories as means to provide climate reconstructions.



Michel is absolutely correct to state that it is not trivial to extend Bclim in such ways, and so we have changed the sentence where we cite PMIP to make this clearer.

So all in all, I would say that the claims for ‘new paradigm’ have to be given more substance. It is necessary to position ‘Bclim’ among recent contributions better, and the role of ‘Bclim’ in this business needs to be more accurately pitched.

As stated above we have only used the terms ‘novel’ or ‘new paradigm’ in relation to that of climate histories.

The authors take great care to expose the principles of Bayesian inversion step by step while avoiding methodological details, using the toy example to explain how probability distributions of climate time-differences cannot be obtained from the ensemble averages. The principle of “inversion” even deserves a Table to insist on the difference between forward modelling and inversion.

We could have argued about the overall benefits and limitations of the Bayesian paradigm as a restricted paradigm of inference in Earth Sciences. I’d be happy to leave that out here, but let’s keep somewhere in our mind that it is not so easy to beat the palynologist’s (or palaeoceanographer’s) intuition about truth with formal statistical methods. Furthermore, the complexity of statistical algorithms form also a layer of epistemic opacity that may be perceived as a problem to some. The authors might want to polish here and there what should be the legitimate pretensions of a Bayesian framework in palaeoclimate but this is not a point I will insist on.

More relevant for this review I believe, the interpretation of posterior priors is arguable. p. 7, it is said that each “set of sample values from the joint posterior distribution is equally likely”, or, “equally probable” as it is stated in the abstract. Strictly speaking the measure of a given history is zero, and all sampled histories will be associated with different values on the posterior distribution function, whether they are near the tails or near a mode of the distribution. This notion of “equally likely” is in fact confused with the notion of equal “weights”, which only takes meaning in the context of the distribution sample.

Michel is, of course, correct in that we should really state that each climate history is equally *unlikely*, having probability measure zero. However, in this paper we have tried to avoid using mathematically precise notation at the expense of clarity, and this seems like a good place to draw the line and avoid confusion. Again, mathematically rigorous readers should return to the RSS paper.

We have not use the word ‘weight’ or ‘equal weight’ anywhere, so it is not clear whether Michel is referring to our confusion or that of the reader. We have not made any changes in response to this.

There is a similar issue p. 13, about the “most representative histories”. Individual histories as such are not “most representative”, but the choice of three histories can be made such as to express the posterior distribution representatively.

We have defined explicitly how we obtain ‘most representative’ histories:

Included in each plot are three most representative climate histories, constructed by finding the histories that are the median Euclidean distance away from the time-wise overall median.

They are ‘representative’ of the posterior distribution in this sense alone.

A much more minor issue : the authors mention ‘three dimensional climates’, which for many reader will be a reference to the 3-D physical space (latitude, longitude, altitude). Some rewording might be useful to avoid confusion.

We now explicitly refer to the three climate dimensions the first time they are mentioned, and point out that this it the definition used throughout the paper.

The demonstration of the case studies is supported by quite good graphics and the main conclusions concern (a) the time-transgressive character of the changes, among the different variables recorded at Roya and Monticchio and between these variables and Greenland, and (b) the greater ‘variability’ of the Holocene compared to the previous interglacial. Point (a) is supposed to be supported by Figures 2 and 3; and (b) by Figures 4 and 5. I note that the tone used to discuss the time-transgressive character remains fairly shy and the supporting physical mechanism is suggested as a possibility only. I had to play with the figures, drawing vertical lines, trying to convince myself,... in a way not unlike what I would have done with climate records that have not received the Bayesian processing presented here. I wonder whether there is any room for alternative graphics, that would focus better on the time of the transition and provide a more convincing demonstration of what the authors want to show.

I have a similar problem with the Eemian / Holocene discussion. Figure 4 shows nothing really convincing about the difference between Eeamian and Holocene

variability, and Figure 5, bottom row, is a challenging one to contemplate. Whether the 'red' (MIS 1) is distinctively different from the 'orange' (MIS 5e) is not obvious to see, and I failed to be wholly convinced by the author claims in the last paragraph of p. 20. Furthermore, there may be some confusion of language, as whether we speak of differences in low-frequency trend (we do indeed expect MIS 5e to vary more than MIS1 over the entire interglacial interval, because of the differences in insolation forcing) or in short-term volatility. As there was some emphasis on volatility in the paper this may be a source of confusion. The reference to stability, p. 21, is thus not accurate enough.

These are all interesting points, though not raised by the other reviewers. Due to the late submission of this review we have not been able to substantially re-work the figures.

p. 4, line 6 : data – > information

Changed

p. 4. 2 lines before section 3. : ... the wider use of recent statistical methods (rather than research findings)

Changed

p. 9., after section 4.: how we taken (misprint ?)

Changed

p. 10. : “The proposed three dimensions of climate are accepted if they match these response surface well” : need to be reformulated or clarified. Do you want to say that the three dimensions are good predictors of the pollen count sample?

This sentence was actually in the wrong place, and has now been removed.

p. 13, paragraph starting with “Figure 2”, l. +4 : “in falls” (and falls ?)

Changed