We thank Luke Skinner for his summaries of the reviews and insights into radiocarbon and benthic d18O age model construction. It is our belief that we can sufficiently address each concern in a revision of the manuscript. Below is a point-by-point response to the editor's comments.

Now that a sufficient number of reviews have been submitted, I would like to invite you provide a point-by-point response to each of the review comments. I would like to underline the importance of addressing a key issue that has been raised in the reviews, namely: the 'empirically constrained' sedimentation rate prior that is applied in the matching algorithm. One issue is that the validity and applicability of this prior, across a range of sedimentary contexts, does not appear to have been fully assessed in a transparent manner in the manuscript – and indeed seems doubtful.

Below I add a few remarks of my own, in case these are helpful for considering what revisions you would undertake, and I invite your response to these too. I find your study of particular interest, and I hope that my comments will be seen as useful.

Title/general ethos:

In general, I think it might be useful to more clearly delineate the distinction between alignment and 'dating' at the core of the manuscript (even though the difference between relative and 'absolute/numerical' ages is indeed noted in the paper a few times). Creating a benthic d18O stack is one thing, aligning to a benthic d18O stack is another, and dating a sediment core is yet another.

As you say, the manuscript addresses the distinction between relative and absolute ages a few times already, but we will attempt to further highlight this in our revision. Your delineation of three separate processes: alignment, stacking, and dating has been true in the past; however, our new approach integrates these three steps together, such that this is no longer the case. Dating from absolute age proxies is now incorporated simultaneously with the determination of relative ages from alignment. This is why we feel it is appropriate to call these "multiproxy age models;" it is not simply the assignment of a radiocarbon age model after alignment or stacking. We will better describe the nuance and novelty of our approach in the revised draft.

The only way that benthic alignment provides age constraints is if one proposes to have prior knowledge of how local/regional deep-water T and d18Osw relate to insolation, e.g. based on a hypothesis for how insolation paces ice volume, and how changes in ice volume are linked to deep water T changes and/or influence deep ocean d18Osw at a given location in the ocean. The latter sequence of hypotheses can give age constraints that are of ~millennial accuracy at best. In such a context, radiocarbon dates (even with ~centennial uncertainties in reservoir age offsets) obviously can provide a refinement.

BIGMACS is not attempting to use the benthic d18O as an indicator of absolute time (e.g., via orbital tuning) as you describe here. d18O is used only to constrain relative ages between cores (or a target stack). We will add more clarification of this point. Our motivation in developing this software is the same point you make above, there is added value in simultaneously considering benthic d18O and 14C during the alignment process.

The inverse is unlikely to be true: age constraints on benthic d18O are unlikely to be precise enough, even to constrain changes in radiocarbon reservoir age offsets of order 100-1000 years.

We do not intend to make such a claim in the manuscript and would welcome information regarding locations of the text that lack clarity regarding this point.

Alternatively, if the core notion of the manuscript and algorithm is the simple transferral of a radiocarbon chronology (or the pooling of radiocarbon dates) between sediment cores via a stratigraphic alignment of benthic d18O, then again it is not quite a case of 'combining age constraints from radiocarbon and benthic d18O'. Rather, it is one of radiocarbon dating of a stratigraphic alignment/stack.

The main concern here appears to be a nuance in how the age model technique is described. Your description of a "transferral of a radiocarbon chronology (or the pooling of radiocarbon dates) between sediment cores via a stratigraphic alignment of benthic d18O" is a reasonable summary of the software during stack construction. However, BIGMACS can also include other types of age constraints, offers an alignment-only mode which does not require any 14C dates, and also constrains age models based on a prior for sedimentation rate variability. While the concept is simple, the probabilistic model leads to rigorous statistical results that simultaneously combines many pieces of information.

What the software does is more complex than "radiocarbon dating of an alignment/stack" because the alignment/stacking is performed simultaneously with consideration of the 14C age information. Because the software integrates relative age information from benthic d18O simultaneously with 14C age constraints, we feel that "multiproxy" is an appropriate description of the age modeling method. In fact, we show in Figure S3 that considering both d18O alignment and 14C dates (ie, multiple proxies) in age model construction produces smaller confidence intervals than either alone (in large part due to the pooling of 14C dates from multiple cores, which requires alignments of the cores).

As such, my own feeling is that the manuscript might more accurately be framed in terms of 'refining orbitally-tuned benthic d18O age models using radiocarbon constraints', e.g. in the title and through the text.

There seems to be some misunderstanding here. No orbitally tuned ages are incorporated in BIGMACS unless a user chooses to align to an orbitally tuned target. None of our examples use any age information derived from orbital tuning. If radiocarbon dates are used in stack construction, the final age model produced will derive almost entirely from 14C over the time period for which 14C dates are available (because the stack is iteratively updated to agree with absolute age constraints).

In a similar vein, it seems to me that describing the age models as 'multi-proxy' is a little misleading: my own expectation was initially of something like that described in line 522. I would again suggest that the process tackled in the present study be described as something like 'radiocarbon-refined single proxy alignment'.

You are likely not the only reader who may initially misinterpret our use of the word multiproxy, and we concede that it is important to clearly explain our use of the term multiproxy in the abstract and early in the introduction (ie, multiple types of age information). Although currently BIGMACS can only align benthic d18O data, we hope to develop a future version will be equipped to handle alignments of multiple climate/sediment proxies.

We refer to the age models as multiproxy because information from both radiocarbon ages and benthic d18O-alignment are simultaneously integrated during age model construction. If the user has an input core with radiocarbon data, benthic d18O data, and an independently dated alignment target that they believe has a synchronous benthic d18O signal, then the combination of both proxies determine the age uncertainty of the age model. Figure S3 demonstrates that the multiproxy age models have smaller uncertainties than their single proxy counterparts for every core in this study.

BIGMACS can also incorporate "multiproxy" (relative and absolute) age information on timescales beyond the limit of 14C dating with the use of additional sources of absolute age information (called "additional ages" in the software). For example, tephra layers or tie points to speleothems can provide additional absolute age information to improve an age model, compared to a "single proxy" alignment based solely on benthic d18O alignment.

Although other age modeling software can combine absolute age information from 14C with non-14C sources, none of the other dating software combines these absolute age estimates with automated signal alignment (ie, probabilistically derived relative ages). For example, Undatable only incorporates relative age information if the user specifies an absolute age estimate for a discrete tie point. This is dramatically different from the continuous probabilistic relative (aligned) age estimates that BIGMACS integrates with absolute age constraints.

Perhaps the source of concern with the term "multiproxy" is that benthic d18O and 14C are proxies for different things. 14C is a proxy for absolute age whereas benthic d18O is a proxy for climate/seawater properties and, thus, relative age. We are happy to add more explanation to clarify our intended meaning of "multiproxy," but we assert that our use of the term is both technically correct and appropriately conveys the integration of multiple sources of (different) age information. However, if a decision about publication of this manuscript ultimately rests on our inclusion of the word multiproxy in its description, we are willing to discuss alternatives.

Line 25:

This line is not quite correct: the accuracy with which ocean sediment cores can reconstruct the *timing* of past climate events, depends on.. the.. age model. The accuracy of proxies is a separate (thorny) matter.

Ok we will revise this sentence to state "The accuracy with which ocean sediment core data can reconstruct the timing of past climate events…"

Line 48:

"Sedimentation rates are realistically constrained…."

As pointed out by the first Reviewer, it seems we must take this on faith, whereas there is burden of demonstration here.

In our revision we will include an in-depth description of the data and construction methods of the transition model. In addition, we will add a section in the discussion that summarizes the strengths and weaknesses of this prior, as well as areas of future improvements.

Line 65:

In general, there is a need to be precise when describing radiocarbon procedures. Radiocarbon dates need to be calibrated to account for past changes in the initial radiocarbon concentration of the fossil entity's 'parent reservoir' (atmosphere, surface ocean, etc.), which may change due to 14C-production changes and/or other carbon cycle processes. This crops up again on Line 72: planktonic foraminiferal radiocarbon dates must be corrected for 'reservoir age offsets' (relative to the atmosphere) only if using a record of past atmospheric radiocarbon concentration/activity for the calibration. In principle, a 'marine calibration curve' might be used instead, with different potential corrections needed as a result.

Thank you for catching this. We will revise these sentences to reflect the processes described in Heaton et al., (2020).

Line 79:

"…requires simulating the core's sedimentation rate."

I think this might be more accurately phrased as: "…requires the assumptions/models of the core's evolving sedimentation rate between dated intervals."

Yes we agree, the sentence will be revised to reflect the above changes.

Line 90:

I think this is a but unfair to Bchron: instead of 'resulting in extreme sedimentation rate variability', it simply posits the full range of possibility wherever there are no prior constraints on sedimentation rates. This is arguably pretty sensible, and it represents a useful counter point to methods that assume a priori knowledge of sedimentation rates.

We can rephrase these sentences to read "Minimal sedimentation rate constraints often result in age models with larger uncertainties than other software packages…"

Line 109:

Again on the sedimentation rate prior issue: does a prior on sedimentation rate not 'beg the question' with regard to down-core changes in age, requiring simply a single point to be anchored in time? This seems like a very (overly) strong constraint to apply, does it not?

I think the confusion here is that the prior describes sedimentation rate *variability*, i.e., it allows for simulation of changes in sedimentation rate that are primarily used to estimate age uncertainties *between* radiocarbon dates or other absolute age estimates. (A prior that specifies a low level of sed rate variability might also reduce the fit to absolute ages if fitting those ages requires large, rapid sed rate variability; thus, the apparent sed rate changes might be ascribed to 14C dating uncertainty.)

In Undatable, the rate at which uncertainty grows between 14C dates is based on a user-specified parameter which is not based on a statistical analysis of sedimentation rate variability in any set of cores. Although the set of cores we use is not comprehensive and future work might improve our prior, the approach of BIGMACS to use a physically based formal prior to describe how age uncertainty varies between radiocarbon dates (and beyond the first and last date) is fundamentally different from that of Undatable.

Line 138:

Is it worth noting perhaps that this shifts the problem of assuming 'instant ocean mixing' to one of a priori knowledge of past ocean hydrography and circulation?

We discuss one possible strategy of assessing homogeneity by using LGM and modern water mass geometry reconstructions in section 4.

Table 1:

note that the 14C dates for MD99-2334K are reported only by Skinner et al., G-cubed 2003 (Skinner & Elderfield 2003 does not exist, and was omitted from the references for this reason no doubt); Skinner and Shackleton 2004, and Skinner et al., Paleoc. & Paleoclim. 2021.

Thank you for catching this, we will correct the citations here.

Figure 6: What is the reason for choosing this sediment core in particular? MD99-2334K is included in the present study, has various alternative stratigraphic age-models (aligned to the Greenland ice core event stratigraphy, and the Hulu speleothem record), as well as a reasonable 14C chronology, and a well resolved benthic d18O record. Would this not be an optimal target for testing the method? A comparison with MD95-2042 could also be made, since both also have 'alignable' planktic d18O records. Furthermore, these two cores were obtained using different coring devices resulting in very different 'apparent sedimentation rates' (due to compaction in the Kasten core and stretching in the Calypso corer), providing a useful basis for assessing the algorithm's sedimentation rate prior.

This is a good suggestion, both MD99-2334K and MD95-2042 would serve as a good example here. We chose GIK13289-2 because it is not included in the DNEA stack (the alignment target) and thus the agreement between the d18O-only and C14-only age models validate our assumption of homogeneity. If we chose MD99-2334K or MD95-2042, we would expect the d18O-only and C14-only age models to agree fairly well because both of these cores contributed to the alignment target.

Line 537: again, I would propose that it might be more transparent to refer to 'radiocarbon-refined/guided d18O alignments, or similar. I wonder what the authors think.

This sentence factually lists the data and prior (sed rate variability model) that the software uses to generate probabilistic age models. It sounds like the main concern here is again the use of the word "multiproxy," which we explained above.

The simultaneous interaction of absolute age information and relative age information, particularly during stacking, is not fully captured by "radiocarbon-refined/guided d18O alignments" because 14C ages simultaneously affect both the individual alignments of the cores in the stack (thus the stack's features) and the stack's age model. "Multiproxy" is intended to convey the integration of multiple sources of (different) age information. If there is any remaining ambiguity in our usage of the word “multiproxy”, we hope for continued correspondence in an effort to make our work as clear and accurate as possible.

I look forward to reading your views on these, and most importantly the reviewers’, comments.

Sincerely

Luke Skinner