Response to Climate of the Past Comments:

On the tuning of plateaus in atmospheric and oceanic ^{14}C records to derive calendar chronologies of deep-sea cores and records of ^{14}C marine reservoir age changes

Edouard Bard[1] and Timothy J. Heaton[2]

1 CEREGE, Aix-Marseille University, CNRS, IRD, INRAE, Collège de France, Technopôle de l'Arbois, Aix-en-Provence, France

2 School of Mathematics and Statistics, University of Sheffield, Sheffield S3 7RH, UK

1st May 2021

We would like to thank all the contributors for their comments on our submission. We are glad to have instigated such an extensive discussion and have found the issues raised very helpful in revising our paper.

In particular, the comments by RC1 (Paula Reimer) led us to perform and provide further statistical calculations to demonstrate the effects of the threshold value of the gradient (^{14}C yr/cal yr), and of the kernel bandwidth used to detect and define the age plateaus. These updated and new figures are included below in the response to RC1 comments.

The comments by RC2 and CC6 (Michel & Siani) convinced us to add a conclusion and outlook section to our paper, underlining the present limitations and emphasizing alternatives and promising ways to improve chronologies and MRA reconstructions (tuning based on paleoclimatic variations, matching floating tree-ring ^{14}C series with ^{10}Be records from ice-cores, MRA records based on the tephra method).

The CC5 (Lamy & Arz) and CC6 (Michel & Siani) provided specific details on cores from the Chile and Brazil margins, reinforcing our initial criticisms of the plateau tuning (PT) applied to these sediments.

The CC1 (Sarnthein & Grootes), CC2-CC3 (Grootes & Sarnthein) and CC4 (Weninger) allowed us to clarify and strengthen our points about the use of PT by Sarnthein et al. (2020) and previous publications.

As underlined by the referees and most commenters, our paper will constitute the first assessment of the PT technique, independent of the Kiel group.

We provide below individual responses to the RCs and to the CCs

Referee Comments (RCs) – Summary of comments and our responses

<u>RC1 – Paula Reimer</u>

We agree that PT needs very strong assumptions as underlined by Referee 1:

_ A constant MRA during plateaus is very unlikely if the cause of the large MRA values reported by Sarnthein et al. (2020) is hypothesized to be due to carbon cycle changes.

_ Predicted abrupt changes in $\Delta^{14}C$ are an order of magnitude larger than any documented production rate changes in the dendrochronologically-dated tree-ring ^{14}C records – no explanation for why this would be the case.

_ Large variation in sediment accumulation rate required to produce the plateaus including hiatuses, but Sarnthein et al. don't provide independent evidence from the marine cores themselves.

RC1: Line 572-3: *'Fig. 5d shows the gradient estimates overlain with the suggested gradient threshold of 0.5 ^{14}C yr/cal yr'. Explain why a gradient of 0.5 was used.*

When estimating the local gradient using the proposed automated approach, no definitive rule was given by Sarnthein et al. (2015) as to either the choice of bandwidth (i.e. how many neighbouring points to use) or the gradient threshold which should be used to define a plateau. The subjectivity in these selections (and the considerable impact they have on plateau identification) reduces the reproducibility and objectivity of the PT technique.

Sarnthein et al. (2015) trialled two threshold values. They state a value of 0 ^{14}C yr /cal yr (i.e. a plateau stricto sensu) generated too many short potential plateau periods. By increasing the threshold value, these disconnected short time periods should merge with one another to create longer time periods. However, other disconnected time periods may be introduced. Sarnthein et al. (2015) therefore also trialled increasing the threshold to 1 ^{14}C yr /cal yr which they suggest agreed better with their visual preferences. However, a time period with a gradient of 1 is quite a distance from what most would describe as a plateau. Indeed, a gradient of 1 is what the slope (^{14}C yr /cal yr) should be without any perturbation of the radioactive decay.

For this reason, we chose 0.5 (^{14}C yr /cal yr) as a gradient threshold predominantly because we felt it was a simple number and we would expect a plateau to have a gradient considerably less than 1.

In response to this reviewer (and also the comments of Grootes & Sarnthein in CC3) we have extended our simulation study to show (attached Fig. 1 – an extension/improvement to Fig. 5 in the original submission) where would have been identified as a ^{14}C-age plateau in each core had a threshold of 0, 0.5 or 1 ^{14}C yr /cal yr been used instead. One can see the choice of threshold makes a considerable difference to the suite of plateaus one might select; further, for any threshold, the plateaus identified are not entirely consistent between cores.

We have also provided a further extension (attached Fig. 2) showing how the estimated local gradient would differ had a larger bandwidth for the kernel been used, i.e. using a wider moving window. This illustrates how different bandwidth choices may significantly affect the resultant ^{14}C-age plateau atmospheric target.

The sensitivity to these choices does not simply create a challenge in identifying individual ^{14}C-age plateaus but also the atmospheric suite of plateaus – both in terms of number and location. Other than the YD plateau, which does seem to be fairly regularly identified, the other periods one might identify as atmospheric plateaus, based on local gradient, do not accurately align between the three cores.

Even if one sought to merge the plateau periods, it is unclear as to how one would objectively know which to merge with others. If one has different numbers of ^{14}C-age plateaus in the atmospheric target suite, then one will tune the marine core very differently. A PT using the blue simulated record as an atmospheric target (panel a) would likely lead to very different results than tuning to the red simulated record (panel b). In particular, we see in panel (d) the lack of consistency in the alignment of the gradient-identified ^{14}C-age plateau periods from 13 – 14 cal kyr BP. This indicates the difficulty of knowing what hypothesized suite of plateaus one should aim to tune when faced with ^{14}C data similar to Lake Suigetsu.

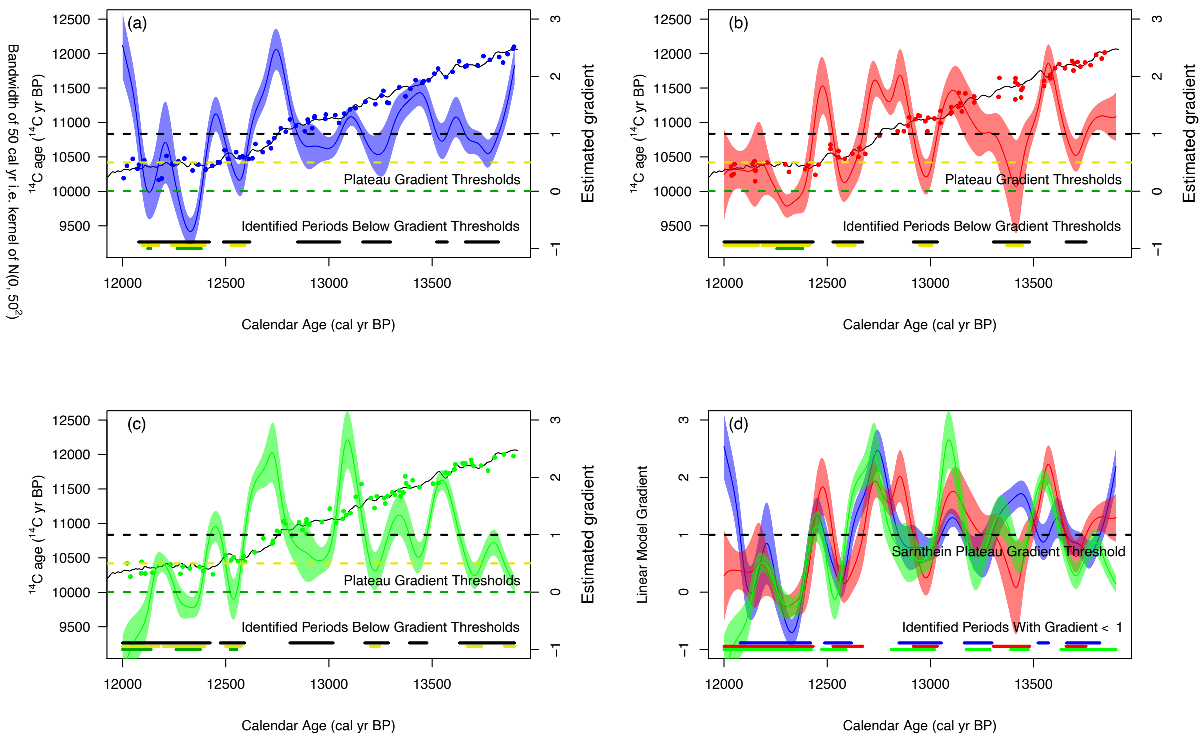

Figure 1 – Update to Figure 5 in original submission. Simulation study to identify the ability of a Suigetsu-style record to reliably identify atmospheric ^{14}C-age plateaus. In panels (a)-(c) we present three simulated atmospheric records generated by sampling, subject to noise, from the high-precision tree-ring-based section of IntCal20 between 12-13.9 cal kyr BP (shown as a black line) with a sampling density matching that of Lake Suigetsu. The level of noise added to create these simulated ^{14}C observations (blue, red and green dots) was also of an equivalent level to that present in the Lake Suigetsu ^{14}C. For each simulated set of observations, we present an estimate of the local gradient (shown as blue, red and green curves with their 95% confidence intervals) according to a locally-weighted linear model as proposed by Sarnthein et al. (2015). These local gradient estimates are obtained using a $N(0, 50^2\ cal\ yr^2)$ kernel to provide the weightings. We overlay three gradient thresholds (0, 0.5 and 1 ^{14}C yr/cal yr) which might be used to identify a ^{14}C-age plateau. Shown as a rug at the bottom of each plot are the time periods in each core which correspond to a local gradient below each threshold (color-coded by threshold). In panel (d) we overlay the gradient estimates to assess consistency (or lack thereof) between the three simulated cores in terms of the number and location of ^{14}C-age plateaus one might identify. As a rug, we plot the time periods in each core (color-coded by core) which correspond to a local gradient below the threshold of 1 ^{14}C yr/cal yr proposed by Sarnthein et al. (2015).

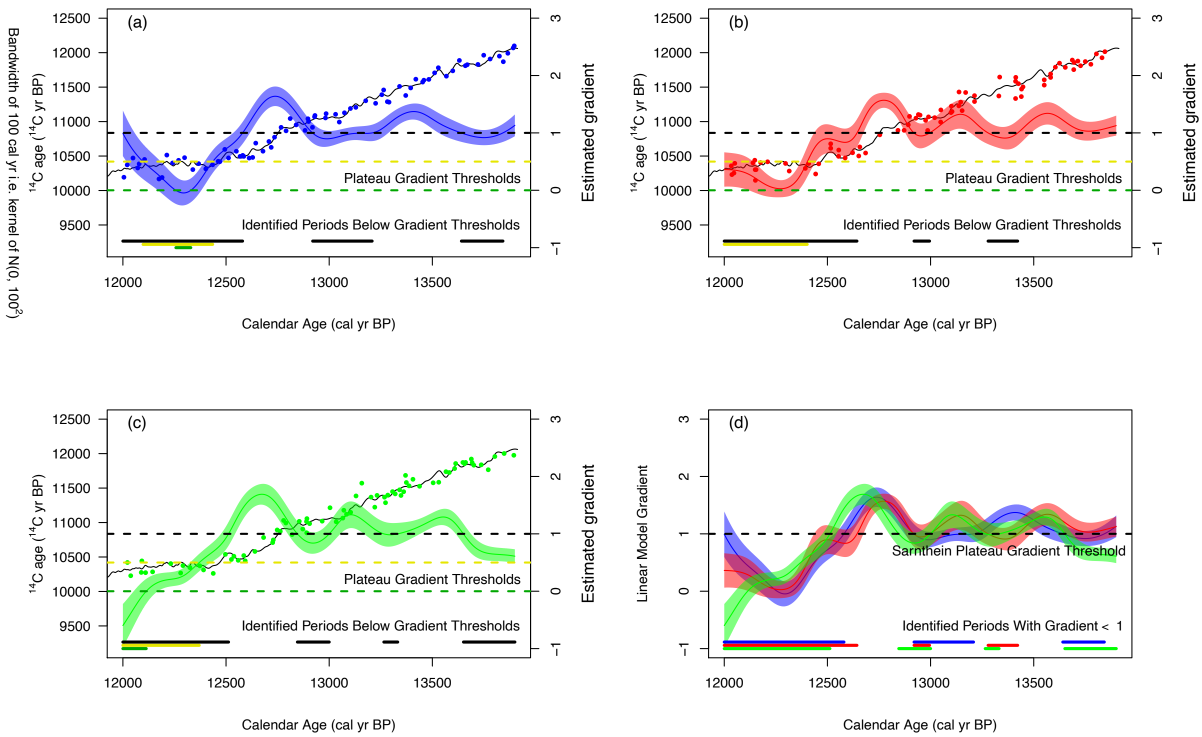

Figure 2 - As Figure 1 but using a wider $N(0, 100^2\ cal\ yr^2)$ kernel to provide the weightings for the local gradient estimate. The same three simulated atmospheric cores are used.

RC1: Figure 6b: *It would be useful to have the gradient threshold marked on this figure to see where potential plateaus might exist.*

We have also modified this figure as suggested, see attached Fig. 3. We have added the same three gradient thresholds of 0, 0.5 or 1 (^{14}C yr /pseudo-cal yr) as for the atmospheric element of our simulation study. These are on a pseudo-calendar scale based upon an assumption of equi-depth spacing of the ^{14}C samples within the marine core (as explained there is no known calendar age scale on which to estimate the ^{14}C-age gradient until PT has already been completed).

In panel b) and c), we have shown the section of the core/time period for each core (colour coded by core) which would be identified as a plateau having a pseudo-gradient < 1 ^{14}C yr /pseudo-cal yr. In panel d) we have shown (as a colour coded rug) the periods one would identify as plateaus in the simulated marine sediment cores (with threshold < 1 ^{14}C yr /pseudo-cal yr) against those in the first simulated atmospheric core (with threshold < 1 ^{14}C yr /cal yr). For PT to be reliable, one would wish these periods to align both in number and location across all the cores. We can see:
1) Beyond the YD ^{14}C-age plateau, the sparsity and the lack of known calendar age scale in marine ^{14}C records make the identification of plateaus very difficult, unreliable and inconsistent,
2) Considerably different numbers and locations of plateaus are indicated using the two simulated marine records; and for all different gradient thresholds,
3) Plot d) highlights the inconsistencies in the periods which would be identified as plateaus, and the errors that would be introduced should one attempt to align the identified plateau periods of either of the simulated marine cores to the simulated atmospheric target.

Again, we note this is not solely an issue of identifying individual plateaus, but rather how one would pair hypothesized plateau-suites when there is such potential ambiguity.

Critically, we wish to emphasize that this simulation study assumes no MRA changes in the marine core – it is simply the effect of sparsity, noise, and the lack of timescale on which to reliably infer the ^{14}C age/cal age gradient (which is inherent to PT). The introduction of MRA changes will add a very considerable further layer of confounding to the identification of any plateaus in the marine sediment ^{14}C record.

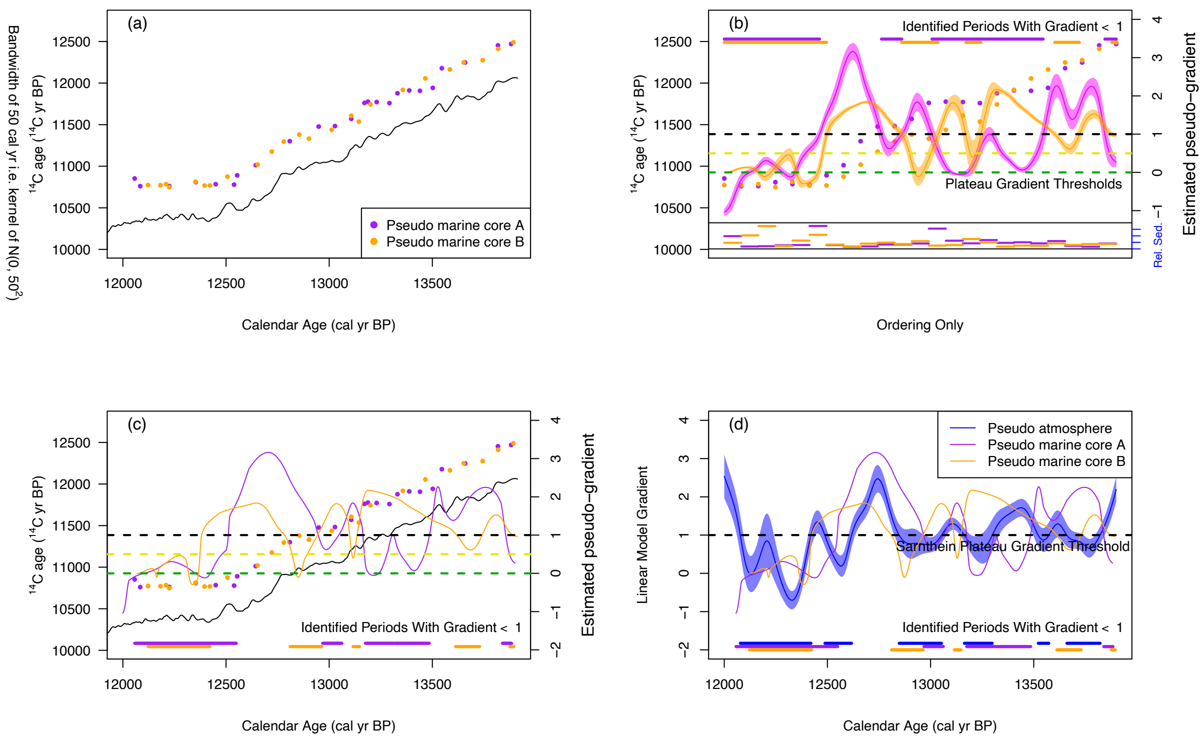

Figure 3 – Update to Fig 6 of original submission. Simulation study to assess the ability to identify ^{14}C-age plateaus in marine records for which calendar age scales are not initially known. In panel (a) we simulate two marine records between 12-13.9 cal kyr BP, again using the IntCal20 tree-ring-based record as the ground truth. These simulated marine cores are based upon the sampling density of the Cariaco Basin unvarved record, with a similar level of observational ^{14}C noise, and have been created with a constant MRA of 400 ^{14}C yrs. To identify plateaus, we first estimate the gradient on the basis of an unknown calendar age scale. We create a pseudo-calendar age scale by rescaling the observations so they are equi-spaced along the cores before applying the same local approach to estimate the gradients on this pseudo (equi-spaced) scale using a $N(0, 50^2$ cal yr$^2)$ kernel to provide the weightings. Estimates of pseudo-gradients, and implied relative sedimentation rates, on this equi-spaced pseudo-timescale are provided in panel (b). At the top, we also plot (color-coded by core) the section of core for which we would obtain a local pseudo-gradient below 1 ^{14}C yr/pseudo-cal yr which could be identified as a ^{14}C-age plateau. Scaling back to the true, underlying calendar age timescale, panel (c) indicates where, in terms of the actual calendar age scale, one might classify plateaus. In panel (d), these are overlain against the first of our simulated atmospheric records from Fig 1 (5 in original submission) to assess synchroneity. As a rug we show the time periods, on the true underlying calendar age timescale, corresponding to local gradient threshold estimates below 1 ^{14}C yr/cal yr for the atmospheric record and 1 ^{14}C yr/pseudo-cal yr for the two simulated marine records.

Other minor comments by RC1 (all addressed as suggested):

Line 141-143: 'Consequently, the total duration of ^{14}C plateaus represent 82% of the time spent between 14 and 29 cal kyr BP, whereas during the remaining 18% of the time, the *radiocarbon clock was running almost 5 times too fast'. This concept needs clarification for most readers to follow.*

The percentage of 82% corresponds to the cumulative duration (≈ 12 kyr) of the plateaus between 29 and 14 cal kyr BP, as defined by Sarnthein et al. (2020), compared to the total duration of this time window (≈ 14.6 cal kyr). During the remaining time (18% ≈ 2.6 cal kyr) the ^{14}C ages change between 24.1 and 12.5 kyr BP (i.e. ≈ 11.6 ^{14}C kyr). Hence, the ^{14}C clock runs 11.6 ^{14}C kyr during 2.6 cal kyr, which gives a ratio of 4.4 (calculating the ratio with the total calendar duration would give a value of 5.6).

We suggest adding clarifying comments to this statement in the paper, "the radiocarbon clock (i.e. the pace at which the radiocarbon age changes compared with true calendar time) was running almost 5 times too fast in comparison to the simple radioactive decay of ^{14}C".

Line 591-2: 'we remove the calendar age information that aids ^{14}C yr/cal yr gradient calculation.' Add that this was done in order to simulate the marine records used by Sarnthein et al. which have no calendar age information.

Paragraph 3.7: Move the following sentence to the start of the paragraph, or re-word, so that the reader knows that IntCal20 is used to simulate the marine record: 'We create our simulated pseudo-marine cores to span 12-13.9 cal kyr BP, again using IntCal20 as our true atmospheric ^{14}C baseline'.

We agree with all these suggestions by RC1 and would be happy to modify our original submission accordingly, including the new figures and associated discussion.

RC2 – Anonymous

We thank this reviewer for the positive comments. We agree with all the points made.

RC2: *authors might consider to add a 'conclusion and outlook' paragraph at the end of the paper, pointing again to the complexity of the link of atmospheric and marine ^{14}C variability in these points:*

We fully agree to add a 'conclusion and outlook section' as suggested when revising our submission, to underline the present limitations and to emphasize alternatives and promising ways to improve chronologies (tuning based on paleoclimatic variations, matching floating tree-ring ^{14}C series with ^{10}Be records from ice-cores, papers that were already cited in our submission). In addition to Waelbroeck et al. (2019) which links to ice cores, we will also discuss the alternative approaches of linking climate proxies within marine sediment cores to oxygen isotopic ($\delta^{18}O$) profile of the U-Th dated Hulu Cave stalagmites (Bard et al. 2013, Heaton et al. 2013, Hughen and Heaton 2020).

Short Comments (SCs) from the scientific community
Summary of comments and our responses

<u>CC4 – Weninger</u>

This comment is predominantly focussed on whether one can identify ^{14}C-age plateaus in the atmospheric record. Importantly, we which to stress that this is only one of the concerns we raise with the PT technique to provide calendar chronologies for deep sea cores. Whether long atmospheric ^{14}C-age plateaus as proposed by Sarnthein et al. (2020), exist or not, this does not imply they can be reliably identified in the marine sediment cores in light of the issues of:
_ varying sedimentation rate within the cores and the consequent lack of calendar timescale on which to estimate the ^{14}C-age to cal-age gradient,
_ the confounding effect of changes of MRA and of bioturbation coupled to foraminifera abundance changes,
_ the typical sparsity and uncertainty of ^{14}C sampling and measurement possible in marine sediment.

The comment includes a Fig. 1 which shows a curve based on an alternative statistical technique (SPD) to detect age plateaus directly in the IntCal20 calibration curve. Dr Weninger claims that all the plateaus named by Sarnthein et al. (2020) can be found in the IntCal20 curve. Beyond the high-resolution part of the calibration, the author suggests his new technique would thus detect the same 11 plateaus between 24 and 14 kyr BP.

For this new work, a recent reference is cited by Weninger & Edinborough (2020) published in *Documenta Praehistorica*, a journal published by the University of Ljubljana. We were not aware of this paper (nor of its journal). In this paper, the authors further claim that the Bayesian statistical methods used for IntCal for two decades, should be replaced by an alternative method, the conclusion of the paper being the following: *"In this paper we propose a rethinking of the mathematical foundation of archaeological ^{14}C-age calibration. We also suggest that archaeologists have at least as much to learn from physicists as they do from mathematicians and statisticians. Following many years of dedicated education, persistent technical support, and admirable instruction by radiocarbon dating experts, parts of the archaeological community are close to the erroneous conclusion that procedures underlying ^{14}C-calibration follow directly from Bayesian probability theory. The choice of a Bayesian framework in ^{14}C-analysis offers, indeed, highly luxurious analytical conditions for archaeological age-modelling. Next to established luxury and acclaimed beauty, the process of ^{14}C-calibration is better described as the Fourier transform."*

It is beyond the scope of our response to evaluate the alternative method proposed by Weninger & Edinborough (2020). However, we can make the following remarks, which relate directly to the subject of the detection of ^{14}C plateaus in the calibration curve:

The curve shown on Fig.1 included in CC4 is the same as the one plotted on Fig. 1 by Weninger & Edinborough (2020). However, in the latter case, the authors only claim the detection of 4 plateaus in the 24 to 14 kyr BP time window of IntCal20, in contrast with the 11 plateaus named by Sarnthein et al. (2020) over the same period. More specifically, the identified plateaus are #1 (which is undisputed), #3, #6b and #8. The Fig. 2 of our submitted paper suggests that plateaus #3, #6b and #8 are dubious in the IntCal20 curve. The same statement is probably true for the Suigetsu record (see Fig. 1 in our original submission), although Weninger & Edinborough (2020) did not apply their technique to the Suigetsu record itself.

Furthermore, there appear calendar time periods in Fig. 1 of Weninger's CC4 comment, for example 18.6 – 18.2 cal kyr and 17.5 – 17 cal kyr, which are not identified as plateaus but to which he assigns equally high shape scores and profiles as other time periods which are classified as plateaus by Sarnthein et al. Again, this relates to the ability to provide a reliable and unambiguous suite of plateaus/atmospheric ^{14}C structure.

We should also clarify that we do not entirely understand the rationale and methodology behind Weninger's proposed approach to identify potential ^{14}C-age plateaus in the IntCal curve. From our reading of the comment, his suggestion appears to be that one would calibrate (against IntCal) a regular grid of hypothetical ^{14}C-age samples (with an assumed common standard deviation) and then sum the distributions of the corresponding calibrated ages. This would appear to be approximately equivalent to providing the solution to the question: *"If one sampled uniformly at random a ^{14}C-age, what is the distribution of the calendar age it corresponds to?"* However, the intuitive solution to this particular question would give a higher summed density to calendar periods where the IntCal mean curve is steep; and conversely a lower summed density to calendar periods where the IntCal mean curve is flat. We therefore assume we have misunderstood. Unfortunately, the paper by Weninger & Edinborough (2020) does not provide the necessary mathematical details to reproduce and test the proposed technique.

However, these concerns are somewhat moot since unfortunately we disagree with the proposed approach which suggests one can easily identify and quantify ^{14}C-age plateaus from the published IntCal mean curve values. Here we agree with Sarnthein & Grootes (CC1 and CC2-3) – that the IntCal mean curve is not intended for this purpose and the smoothness, or otherwise, of these pointwise mean values must be used cautiously when inferring the level of atmospheric ^{14}C variation.

To attempt to use the published IntCal mean values in the way proposed by Weninger (CC4) is to somewhat misinterpret what the calibration curve represents – which is to provide "optimal" pointwise estimates, specifically the posterior mean (with corresponding variance) given the observed ^{14}C data which goes into curve construction. Where one has low sampling density, or uncertain calendar ages for the ^{14}C data, the flatness or otherwise of the posterior mean alone should not be used to indicate the presence or absence of a ^{14}C-age plateau.

One can perhaps see the issue most clearly with two thought experiments. Firstly, suppose one has a calendar time period which has no ^{14}C data on which to estimate the calibration curve. In such a time period, due to the absence of data, we will typically estimate a posterior mean calibration curve with a ^{14}C-age to cal-age gradient of approximately 1. This is our prior, that ^{14}C acts as a reliable geochronometer and, without any data, we have no reason to assume it to be incorrect in the time period under consideration. However, it is quite possible that there may be a ^{14}C-age variation (or even a plateau) in this period, simply that we have not collected the data to observe it. However, were we to simply consider the gradient of the posterior mean curve we would assume no such variation or plateau would exist.

This *lack-of-information* effect can be seen in a practical setting when comparing the sections of IntCal20 around 13.2 – 11.8 cal kyr BP (for which highly-resolved tree-ring ^{14}C data have only recently become available) against the same periods in IntCal13 (which was based on less highly-resolved and sparser data). Due to the lack of data, the published IntCal13 mean is smoother and lacks several short ^{14}C-age plateaus and variations. These short periods of structure were only able to be identified in the IntCal20 mean due to the increased availability of data. The IntCal13 estimates are not wrong *per se* – after adjusting for the calendar-age correction affecting this portion of the curve described in Reimer et al. 2020, the IntCal20 estimates generally lie within the probability bands of the IntCal13. However, due to the lack of information available in 2013, one cannot identify these genuine plateaus using simply the IntCal13 mean.

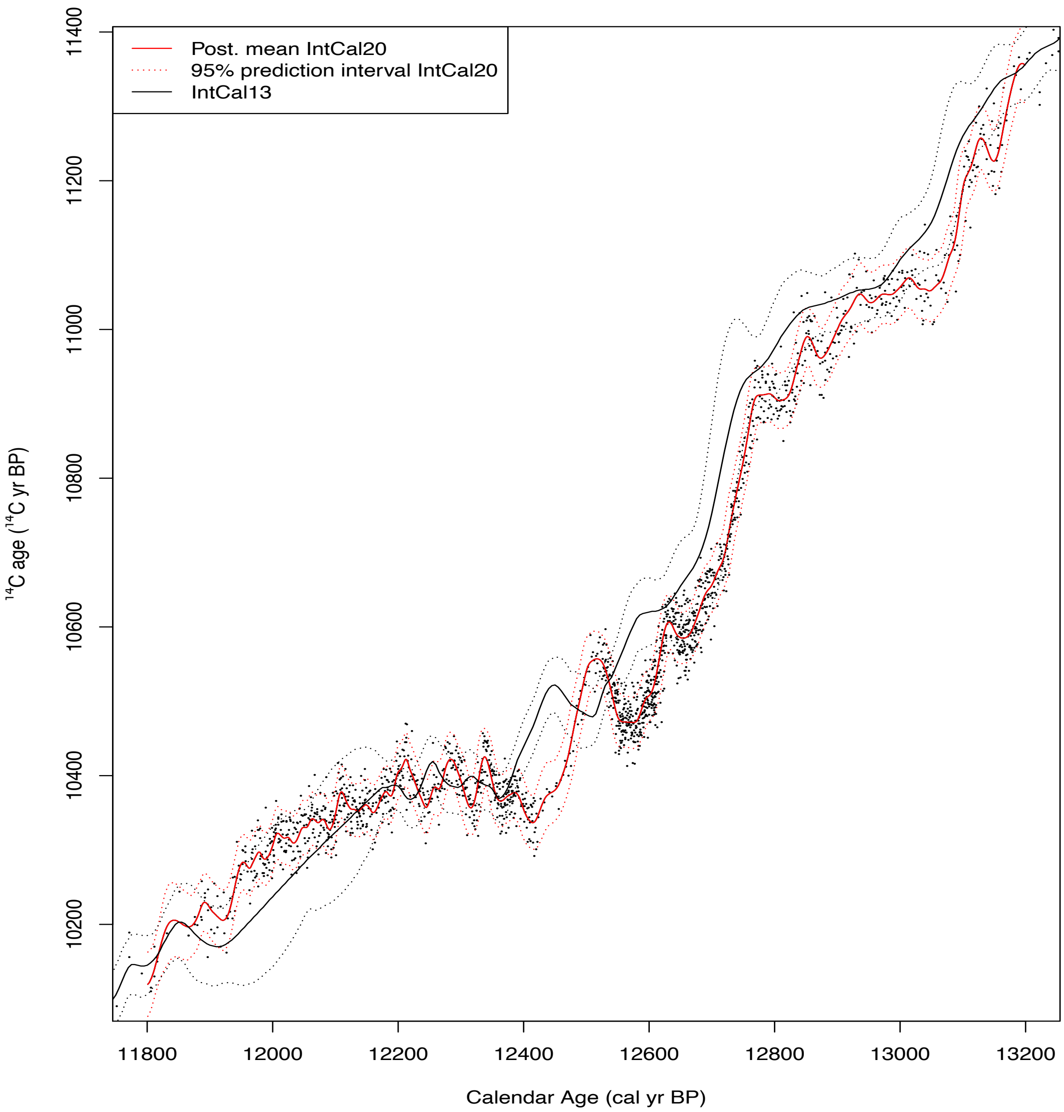

Figure 4 - A plot of the detailed and variable IntCal20 mean vs the smoother IntCal13 mean from 13.3 – 11.8 cal kyr BP. For the IntCal20 curve, high-resolution tree-ring ^{14}C measurements became available covering this time period. This has revealed significant additional structure in the IntCal20 curve which was not represented in the smoother IntCal13 mean. Note however that, after adjusting for the calendar-age shift which also occurred in the underlying tree-ring chronologies between IntCal13 and IntCal20, the IntCal13 probability envelopes generally encapsulate the IntCal20 mean. The IntCal13 values are not therefore incorrect, simply lacking in detail and smoother as the underlying data was not available. However, the additional detail discovered in IntCal20 indicates the danger in over-interpreting the smoothness of the IntCal mean when underlying data is lacking – as is the case for much of the period from 55 – 14 cal kyr BP.

Secondly, consider the hypothetical situation where we have a floating tree-ring sequence which does show a short ^{14}C-age plateau bounded by sharp increases. Since the sequence is floating, we do not know precisely when this internal plateau occurs. In creating the calibration curve, we have to average over all the possible calendar ages of this tree ring sequence. While for each possible calendar age for the start of the floating sequence there will be a ^{14}C-age plateau, since we do not know which is the correct starting age, once we average over them our best posterior mean ^{14}C-age estimate at any specific calendar age will simply be the average of the entire ^{14}C tree-ring sequence. Consequently, we will observe a ^{14}C-age plateau in the *mean curve* that extends over all possible calendar ages of the sequence. This *mean curve* plateau will be much longer than the short ^{14}C-age plateau embedded within the tree-ring sequence for which we have evidence.

Using the pointwise posterior IntCal mean to identify ^{14}C-age plateaus in the finely resolved tree-ring section of the calibration curve (from 14 – 0 cal kyr BP) is likely reliable due to the high density of the underlying data. However, using the posterior pointwise IntCal mean, without also carefully considering the curve's uncertainty envelope, to identify potential ^{14}C-age plateaus in the older section from 55 – 14 cal kyr is not.

CC5 – Lamy and Arz

We thank these authors for their personal insights based on extensive work on marine sediments, notably on cores from the Chile and Brazil margins also used by Sarnthein et al. (2020).

In particular, they confirm the serious problem linked to the application of PT to core PS97/137, explaining why its MRA record has been changed drastically between the submitted and published version of the paper by Sarnthein et al. (2020), as we noted in our submission. We note their comments regarding the highly variable sedimentation rates within this deep-sea core, which make the identification of any ^{14}C plateaus, should they exist, highly uncertain.

They extend this comment to other sediment cores studied by Sarnthein et al. (e.g. GeoB3910 off Brazil). Also, they express their doubts concerning the likelihood of hiatuses, which Sarnthein et al. suggest are not only frequent occurrences in sediment cores but more common in cores for which the sedimentation rate is otherwise high.

Finally, they also state that the status of the laminations in core PS97/137 is still debated, which implies that a rough count cannot be used to support the chronology of that core based on PT, as advocated by Sarnthein et al.

All these concerns by Drs. Lamy and Arz, based on their first-hand experience on these cores, agree with the concerns expressed in our paper.

CC6 – Michel and Siani

We also thank these authors for their insightful and positive comments, with which we agree. The precise calendar dating of sediment cores with varying and unknown MRA is indeed a challenging problem, and any technique which is able to do so would be most welcome.

The authors note that Sarnthein et al. (2020) applied PT to core MD07-3088, which they have studied previously (Siani et al. 2013). They express doubts in the reconstructed variability of sedimentation rate based on PT (up to a factor of 25 for that core) and underline that the MRA cannot be precisely defined for the glacial part of the core.

Further we agree that the use of common horizons such as volcanic ash shards (tephra) provides a rigorous approach for reconstructing MRA in the past, even if finding such tephra is non-trivial and may only be sparsely available in specific regions. As cited by the authors, the use of tephra for reconstructing MRA was introduced by Bard (1988) and first tested by Bard et al. (1994). We thus fully agree that PT should be tested systematically and thoroughly with independent MRA estimations based on the reliable tephra method.

In addition, the chronologies based on PT could also be compared with those obtained independently with tuning of paleoclimatic records in marine sediments with those well dated in polar ice cores or speleothems. We also agree with the note of caution that tuning via climate proxies relies upon an assumption of global synchronicity in the climate changes one intends to tune – although evidence does support an assumption of globally-synchronous timing for certain rapid paleoclimatic changes (Corrick et al. 2020).

We have aimed to incorporate these comments in our 'conclusion and outlook' section alongside those of RC2.

CC1 – Sarnthein & Grootes
CC2-CC3 – Grootes & Sarnthein

We thank the authors of these comments, as they are the proponents of PT and authors of the paper by Sarnthein et al. (2020) which is the subject of our paper submitted as an extended comment.

The CP editorial office refers to CCs as short comments (SCs) from the scientific community. Hence, we were surprised to read a total of about 50 pages by Sarnthein & Grootes and Grootes & Sarnthein (CCs and their supplements), usually reiterating their opinions already expressed in Sarnthein et al. (2020). Nevertheless, we respectfully understand that these authors disagree with our views on PT and our criticisms of their paper.

In the following response, we will concentrate on the points for which they bring new testable information, or on the issues that have been misunderstood or misrepresented from our critical comments. The points by Sarnthein & Grootes (SG) and Grootes & Sarnthein (GS) are grouped thematically.

About sedimentation rate variations

Page 1 (SG): We show proof that results of PT of marine sediment records are hardly affected by bioturbational mixing and changes in foraminifera abundance, given the limitation of PT to cores with sedimentation rates >10 cm/kyr;

Page 7 (SG): In contrast to claims of B&H, most short-term changes in sedimentation rate between consecutive plateaus are low, hardly exceeding a factor of 1.5-2.0.

Page 9: (SG) -- We only applied PT to cores with average sedimentation rates of >10 cm/kyr. In many cores the rates exceed 20-40 cm/kyr and go up to >200 cm/kyr.

Page 10 (GS): a major problem appears to be the fluctuating sedimentation rates that result from the introduction of multiple time markers, i.e. plateau boundaries, in the sediment records. Replacing 'constant sedimentation rates' that result from a lack of data with a pattern of significantly varying sedimentation rates can be disturbing but may also provide rewarding insights.

We are unclear why the authors persist in claiming that they apply PT only to cores with sedimentation rates > 10 cm/kyr and where changes in sedimentation rate between plateaus are typically small. In fact, the output of their PT is usually a highly variable sedimentation rate within a single core, by up to a factor of 5 to 8. For example:

_ Core PS2644 (Sarnthein et al. 2015) has an inferred sedimentation rate generally below 10 cm/kyr from 21 – 17.5 cal kyr, reaching as low as 3.8 cm/kyr for a duration of about 1000 cal years. The overall sedimentation rate varies by a factor of 5 within this core.
_ Core MD08-3180 (Sarnthein et al. 2015) also has an overall sedimentation rate that varies by a factor of 5.
_ Core ODP1002 (Sarnthein et al. 2015) has frequent abrupt changes in PT sedimentation rates over plateau boundaries around (and exceeding) a factor of 2.5. For example, from 130 cm/kyr down to 49 cm/kyr; and then from 42 cm/kyr up to 98 cm/kyr and then back down to 36 cm/kyr. Similar scale short term sedimentation rate changes are also seen in ODP 893A (Sarnthein et al. 2015).
_ Core PS75/104-1 (Küssner et al. 2018) has PT inferred sedimentation rates changing from 21 cm/kyr down to 9 cm/kyr and then further to 3 cm/kyr in the course of only a few hundred calendar years beginning around 14 cal kyr from the end of plateau #1 to the beginning of #1a; core KNR-159-5-36 GGC (Küssner et al. 2018) has sedimentation rates changing from 14 cm/kyr down to 6 cm/kyr and then back up to 19 cm/kyr in the time period from plateau #1a to the YD.

In other cores from the Nordic Seas (Sarnthein & Werner 2017), the variation in inferred sedimentation rate within a single PT core is even greater, by orders of magnitude. In practice, the variations are even more extreme since PT frequently implies the presence of hiatuses (sedimentation rate equals to 0). It is indeed these extreme variations of the inferred sedimentation, which should be tested with an independent technique in order to prove that PT is reliable and that sedimentation rate variations are not an artefact of PT. We also note that the sedimentation rate variability implied by PT is also a major concern expressed by the other commenters (Lamy & Arz, Michel & Siani, see sections above).

A focus on the average sedimentation rate over a complete core (proposed as a PT criterion by CC1) is not sufficient. Indeed, it is the profile and range of sedimentation rate within a sediment core which is most critical. This also applies to the impact of bioturbation since the smoothing and phasing effects are directly related to the ratio between the bioturbation depth and the sedimentation rate (this ratio being the average residence time of foraminifera in the bioturbation zone, e.g. Bard et al. 1987).

Page 8 (SG): Though widely not appreciated by paleoceanographers, hiatuses appear to be a feature actually widespread at high-sedimentation rate sites in the deep sea – One may assume: The higher the rates the more extreme they may be subject to changes in depositional regime.

We think that this statement is hypothetical and counterintuitive as also underlined by Lamy & Arz (CC6). What is the evidence for such a positive correlation between hiatus frequency and sedimentation rate, other than through PT? The extreme variations of the PT inferred sedimentation, including frequent hiatuses, should be tested by independent techniques in order to prove that PT is reliable and that sedimentation rate variations and hiatuses are not artefacts of PT. For these crucial tests, PT should be performed completely independently from the results obtained with other techniques (e.g. tuning with tephra or with climate proxy records). In other words, these other time markers should not be combined to PT if one wants to test the validity of this method.

About the assumption of constant MRA during age plateaus

Page 4 (SG): Most ^{14}C plateaus cover time spans of 300-700 yr each, rarely reaching up to 1100 yr, in agreement with Fig. 3 of B&H. We see no problem in accepting that the ocean carbon cycle and MRA have in most cases not been subject to major changes over these time spans and that changes were confined to short intervals in between.

Page 10 (GS): The pattern of horizontal and sloping line segments in the ^{14}C age/cal age domain facilitates visual comparison but has as a consequence that MRA values, obtained from the difference between oceanic and atmospheric ^{14}C plateaus, are constant over a plateau and change in steps from one plateau to another.

We thank the authors for these clear statements about the inherent assumption of constant MRA during ^{14}C-age plateaus which is central in our criticisms (sections 2.2, 2.5 in our original submission).

Contrary to the statement above, the inferred plateaus are indeed longer than 300-700 yr as their average duration is 800 yr over the period of interest (14-29 kyr BP, cf. our Figs. 1 & 2). As underlined in our paper, the cumulative duration of age plateaus is 82%, which implies that MRA could have varied only during 18% of the glacial period. Hence, PT relies on questionable inherent assumptions about when MRA and sedimentation rates changed in apparent synchrony during a very limited fraction of the past (18%).

Page 5 (GS): There is no assumption for MRA estimates! MRA = (pla – Atm) for each plateau pair.

The statement about no MRA assumptions being needed for PT appears to be only valid if one already knows where the atmospheric ^{14}C-age plateau are within the marine sediment.

However, one does not. An MRA change may make an atmospheric ^{14}C-age plateau appear as a non-plateau in a marine ^{14}C record or vice-versa. This will hinder the identification of the atmospheric ^{14}C-age plateaus within the marine sediment core, affecting in turn the MRA reconstruction. Implicitly the PT method must assume that MRAs can only change at plateau boundaries, remaining constant during the course of each atmospheric ^{14}C-age plateau.

About our box model simulations

Page 2 (SG): The box model discussion is scientifically correct. However, it only deals with Pla, the planktic ^{14}C concentration of ocean surface waters, and not with MRA = (Pla-Atm).

Page 11 (SG): The discussion of B&H focuses on changes in the surface ocean and correctly describes the limitations of ^{14}C variability in this reservoir. Yet they forget that MRA = (Pla - Atm) and that the large variability of atmospheric ^{14}C, as seen in Miyake events and the bomb spike, means that also MRA can show variations much larger and more rapid than displayed by the box model.

Page 3 (GS): Their modelling addresses, however, only one facet of MRA, and a strongly attenuated pla signal will generate an MRA (= pla-Atm) signal with little attenuation. This is borne out by the effects on MRA of the ^{14}C bomb spike and Miyake events mentioned in their text. A box model, moreover, does not consider local variations in near- surface ocean mixing and ocean-atmosphere exchange, that can lead locally to large and rapid changes in pla and thus MRA for an unchanged atmosphere.

Page 11 (GS): The modelling and discussion of lines 284-336 of B&H focus on the 14C concentrations of surface ocean reservoirs (pla) of complete ocean basins each in response to an atmospheric ^{14}C production signal. The 12-box model of Bard et al., 1997, introduced in lines 284-294, calculated the changes in ^{14}C distribution over the various reservoirs of the carbon cycle in response to a reduction of the global thermohaline circulation from today's 20 Sverdrup (Sv) to a postulated glacial 10 Sv (Fig. 4, Δ^{14}C values / Δ^{14}C in brackets). In both cases the atmosphere is the benchmark and its Δ^{14}C has been set to zero. In reality the Δ^{14}C of the 10 Sv atmosphere had increased by 35 ‰ relative to the standard atmosphere of 20 Sv.

B&H do not consider MRA, the difference between troposphere and surface ocean in Fig. 4b, only pla. For a large change in Atm, due to ^{14}C production or remote ocean outgassing, the calculated local change in pla may be small, but the change in MRA may be large and, for strong attenuation, approach in shape and size the Atm signal. Fig. 4 thus does not prove that large and rapid MRA signals are physically unrealistic.

Our box-model simulations show the attenuation of an atmospheric ^{14}C change corresponding to an age plateau for two surface ocean reservoirs characterized by very different reservoir ages (300 and 900 yr). The main goal is to demonstrate that these age plateaus almost **disappear** and are delayed in the surface ocean due to the smoothing effect of the carbon cycle, even if it stays strictly constant as assumed for these simulations.

PT relies on the assumption that the plateaus have exactly the same shape, the same duration and the same timing (in cal yr) in the atmosphere and ocean, which is clearly not the case in our simulations. We trust that similar results would be obtained with more complex models forced by the same atmospheric signal. A prerequisite of the MRA reconstruction by PT is the proper identification and tuning of marine and atmospheric age plateaus, which is shown to be difficult or even impossible in the simulated cases.

Our box-model simulations are not intended to simulate ^{14}C and MRA changes due to variations of the ocean circulation. It is thus unclear why the authors cite our work on other simulations based on the same model in which the thermohaline circulation was reduced by a factor two, which obviously impacted all carbon cycle reservoirs.

Actually, we do refer to these other simulations in our section 2.3 where we explain that it is highly improbable that the atmospheric ^{14}C record is essentially characterized by ^{14}C age plateaus separated by instantaneous drops of ^{14}C ages (Figs. 1 & 2 in our original submission). In the $\Delta^{14}C$ space, these instantaneous ^{14}C age drops correspond to instantaneous rises of atmospheric $\Delta^{14}C$ ranging from 50 to 250 ‰ (Fig. 3 in our original submission).

Indeed, part of the simulations presented by Goslar et al. (1995) were obtained with the very same box-model used for our simulations. This was explained in section 3 of our submission about these simulations, complemented by other model studies from the literature: "Finally, it is unlikely that abrupt changes of the carbon cycle are responsible for such large, frequent and very abrupt $\Delta^{14}C$ spikes. For example, switching down the deep ocean circulation instantaneously in a carbon cycle box-model, leads to a rather slow and limited $\Delta^{14}C$ rise in the atmosphere over several centuries (e.g. see Fig. 4b by Goslar et al. 1995 or Fig. 5 by Hughen et al. 1998; see also simulations performed with more complex models by Marchal et al. 2001, Delaygue et al. 2003, Ritz et al. 2008, Singarayer et al. 2008). Consequently, the PT underlying assumption that the radiocarbon calibration curve has the shape of a staircase, is in conflict with our basic understanding of ^{14}C as a tracer."

About foraminifera abundance changes

Page 8 (GS): -- PT of Core SHAK6K-05 provides a nice test case (Fig. 1) to compare short-term changes in the abundance of the planktic foraminifer Globigerina bulloides with the position and length of paired ^{14}C plateaus (Ausin et al., 2019 and 2021). In contrast to conjectures of B&H, none of the twelve plateaus up to >15 cm long is linked to any abrupt change in species abundance.

In our section 2.6 we stated that ^{14}C-age plateaus in marine sediments can also be linked to bioturbation coupled with foraminifera abundance changes (Bard et al. 1987, Costa et al. 2017). For this reason, we recommended to show the absolute abundance records of the foraminifera species used for ^{14}C dating, which has never been the case in papers based on PT by Sarnthein & Grootes.

These authors now cite a new paper by Ausin, Sarnthein and Haghipour (2021) based on a sediment core from the Iberian margin for which they provide the foraminifera abundance counts, showing no obvious correlation between ^{14}C plateaus and drops in abundances. We congratulate the authors for this inclusion, even if these counts are unfortunately not shown, nor provided in the paper.

This being said, we also express a note of caution about this new work. Indeed, Ausin et al. (2021) obtain low MRA and benthic ^{14}C values for that core, notably during Heinrich stadial 1 and the LGM, in stark contrast with records obtained on nearby cores (Skinner et al. 2014, 2021). The MRA drop (down to 300 yr) during HS1 based on PT is also in conflict with modeling results (Delaygue et al. 2003, Ritz et al. 2008, Franke et al. 2008, Butzin et al. 2017),

Although Sarnthein & Grootes present the new paper by Ausin et al. (2021) as a "nice test case" of PT, the strong disagreement with the literature is not reassuring. Furthermore, it also remains to be seen if the bioturbation/abundance couple could not be an explanation for some age plateaus in the 20 other published records based on PT and for which the foraminifera counts are not available.

About specific records

Page 4 (SG): Examples of independent PT confirmation are:
-- PS97-137 off Southern Chile (Küssner et al., 2020): A rough count of sediment laminations has fairly well confirmed the length of a PT-derived paired ^{14}C plateau for the LGM.
-- MD07-3088 off Central Chile (Küssner et al., 2020): Sedimentation rates and ages are confirmed by succession of four independent age values of ash layers.

As underlined in CC5 by Lamy & Arz, the rough counts of sediment laminations cannot be used to confirm the PT tuning in core PS97/137. Criticisms are also expressed in CC6 by Michel & Siani concerning the tuning of core MD07-3088, which they studied previously. See the sections above concerning these two CCs.

Page 8 (SG): B&H are concerned about recent changes in our plateau assignment for two South Pacific cores. These changes are the result of a valuable discussion on alternative tuning modes ongoing after a first public display of data in CPD. Finally, we choose the mode better supported by various lines of sediment-based evidence.

Page 12 (SG): In part, we handled the problem by frankly discussing alternative tuning modes (e.g., Küssner et al., 2020). In part, we admitted minor refinements in the mode of tuning of a suite of plateaus in papers published later-on, that is, as soon as additional lines of independent evidence were available.

In our paper we spotted discrepancies between the MRA results presented for four cores in the submitted and published versions of the paper by Sarnthein et al. (2020). In the two versions, these new results were referred to an unpublished paper by Küssner et al. submitted to *Paleoceanography & Paleoclimatology*.

We note that Drs Lamy, Michel and Siani were coauthors of the version referred to in the submitted version of Sarnthein et al. (2020), while Dr Lamy is no longer a coauthor of the same paper cited in the published version of Sarnthein et al. (2020). In any case, the paper by Küssner et al. is still unpublished today so no explanation is available to the reader to assess the reason for the drastic change of MRA reconstructions between the two versions.

In their CC5, Lamy & Arz have expressed their doubts about the identification of ^{14}C plateaus in these south Pacific cores, notably PS97/137-1, leading to highly uncertain PT tuning with highly variable sedimentation rates.

In their CC, Sarnthein & Grootes state that these hesitations are the results of normal scientific discussion. We think it is not reassuring for the validity of PT if MRA results can be changed so drastically without notice. This suggests that PT is highly subjective and non-robust. We note that these issues are only known to the reader because of the open review policy adopted by *Climate of the Past,* providing in open access the submission, discussions and published version.

Page 10 (GS): The reference in lines 64-65, and 212-224 to Umling and Thunell., 2017, using PT refers to 'puzzling results' but fails to mention the close agreement between the ages they obtained by three independent techniques, including PT, and shown in their Fig. 4. Their Fig. 3c suggests a naming problem for the youngest, 12.8-13.1 cal kyr BP, of their suite of five matched plateaus. This is the 'no name' plateau of SA2020, instead of the 'YD' plateau, which reduces the hiatus from 1200 to ~600 years. Its timing roughly corresponds to the Inter- Allerød-Cold-Period in Greenland/Northwest Europe.

Indeed, we find the Fig. 3c by Umling & Thunell (2017) particularly unconvincing as an example of PT application. In addition, we underlined that deglacial ^{14}C reservoir ages reconstructed for that core exhibit discrepancies with the nearby record obtained by de la Fuente et al. (2015) on another core from the Eastern Equatorial Pacific collected at a similar depth (2.9 km). We note that in their CC, Sarnthein & Grootes propose an alternative PT interpretation than the one chosen by Umling & Thunell, leading to a reduction of a puzzling hiatus in the middle of the record.

About detecting the plateaus

Page 8 (SG): We agree with B&H that aligning the entire ^{14}C record of a marine sediment core with that of the Suigetsu target curve, analogous to the wiggle matching technique for tree ring sequences, is the approach of our PT technique. Since our first paper of 2007 we stress the need that ^{14}C records should be aligned as a whole with their shape, not just with piecewise constant or slightly different offsets within and between the plateaus, the key to our MRA estimates.

Page 4 (GS): Line 40-41: A small but crucial element is missing in the introduction of PT. The technique tunes a suite of ^{14}C age plateaus in a sediment record against a suite of plateaus defined in the atmosphere (as represented by the Suigetsu record)

In our submission, we did not intend to give the impression we believed PT is restricted to the tuning of single plateaus. We recognize that a suite of ^{14}C age-plateaus is identified and tuned. However, one cannot reliably PT if one cannot reliably identify and define an individual ^{14}C plateau. A suite of plateaus does not necessarily add strength – in fact it potentially makes it more challenging should one misidentify plateaus in either the atmospheric target or the sediment record.

The reliable identification and use of a suite of ^{14}C-age plateaus cannot be separated from the identification of an individual ^{14}C-age plateau. Firstly, the authors present complete core chronologies suggesting one does require identification of all the hypothesized plateaus, PT introduces hiatuses where they fail to identify plateaus. Secondly, if one is unable to identify all individual ^{14}C-age plateaus in either the atmosphere or the marine sediment, how can one decide which ^{14}C-age plateaus to align/tune to one another – what is the right match? Further, can one be confident that those plateaus one has identified are correct?

This task is even more difficult in the presence of secondary plateaus linked to artefacts in the marine sediments (e.g. linked to abrupt change of sedimentation or to the bioturbation/abundance couple). The exclusion of these secondary marine plateaus is not straightforward because they could be correlated to a true atmospheric plateau by assuming a change of MRA. PT could thus lead to spurious results, with compensating errors on the calendar age and on the MRA reconstructions. The number of unknowns being large, PT may lead to non-unique solutions.

In addition, while perhaps having some conceptual similarities, PT of marine sediment cores is not equivalent to the wiggle-matching technique to date tree-ring sequences. Wiggle matching relies upon a known internal chronology within the tree-ring sequence – the only unknown when wiggle-matching is the start date of the sequence. Further, when wiggle-matching a tree-ring ^{14}C sequence there is no variable offset to the atmospheric ^{14}C record, and all ^{14}C determinations are used in the matching. This is not the case for PT.

We cannot wiggle-match a marine core to an atmospheric ^{14}C record unless we both:
_ Have a known internal chronology in the marine core (or make strong *a priori* assumptions about changes in sedimentation rate);
_ Already know the MRA offset/depletion between the marine and atmospheric record (or again make strong *a priori* assumptions about its nature both in terms of value and changes over time).
These assumptions run contrary to the proposed aim of PT which is to provide a complete chronology for the marine sediment core without making *a priori* assumptions about either the sedimentation rate or the MRA. Additionally, in PT one only aims to match single boundaries as opposed to utilising the entire ^{14}C record.

Should one know the MRA in a marine core, it would be possible to adapt wiggle-matching to provide a calendar age chronology by stretching and squashing the sediment core to agree with the atmospheric record in light of the known offset. However, when the MRA is not known, it is unclear how wiggle matching could be rigorously applied.

Page 8 (GS): The following sentence 'They are not replicated in either our statistically-robust Lake Suigetsu-only curve (Fig. 1) or in the IntCal20 curve (Fig.2).' is, however, not supported by Fig. 1. Inspecting the green plateaus, selected from the noisy Suigetsu data set by visual inspection as well as by a first-derivative kernel technique, shows that 13 of the plateaus correspond with calendar age broadening in the Bayesian-spline- generated pink band; solely plateaus 5b and 10b are not reflected in the pink band, although the local scatter of Suigetsu data makes their selection understandable.

Page 15 (GS): B&H fig. 1 and 3a actually show a convincing match between the Bayesian-spline generated Suigetsu curves and our selected plateaus. The argument that fluctuations in the section of IntCal20 considered could represent random scatter, i.e. the null hypothesis, can statistically not be rejected with 95 % probability. This means there is indeed no statistical proof that the observed fluctuations in the record are real. Yet, absence of statistical proof is not a statistical proof of absence.

We agree entirely that, when it comes to the variability, or otherwise, of atmospheric ^{14}C levels from 55 – 14 kyr cal BP that "absence of evidence is no evidence of absence" – we have aimed to make this clear both in our submission and response to Weninger (CC4). Indeed, we expect atmospheric ^{14}C variations that will be discovered as new ^{14}C archives (such as new subfossil trees and macrofossils similar to those found in Lake Suigetsu) are recovered from this time period.

However, it is important to stress that an argument that there must be ^{14}C variation in this period, and consequently that PT is valid is something of a *straw-man*. The presence (or absence) of atmospheric ^{14}C structure, even if it has been correctly identified, is not sufficient for PT to be reliable.

It is correct that the published IntCal20 mean does not exhibit the expected, but currently unknown, short term ^{14}C variation from 55 – 14 kyr cal BP. However, this is intended and due to the design of the IntCal synthesis which aims to provide pointwise mean (and variance) estimates for atmospheric ^{14}C levels. Where we lack information, the IntCal method will produce a smooth posterior mean. Any argument about the lack of structure in IntCal20 being "*far more speculative and dubious*" is incorrect – the IntCal method simply does not wish to create potentially spurious structures when there is no evidence for them. As we showed in Fig 7 of our submission and explain in our response to Weninger (CC4), one should not assume that the published IntCal pointwise mean possesses the same variability as the atmosphere, or use it to infer variability without significant care.

One of our concerns is indeed that currently we cannot reliably identify finer-scale atmospheric ^{14}C structure from the Lake Suigetsu record, and that evidence for the hypothesized suite of ^{14}C-age plateaus proposed by Sarnthein et al. (2020) is therefore weak. However, this is only one element of the concerns we present and will simply compound the effect of other more fundamental and intractable issues.

We believe it is dangerous to over-interpret high frequency fluctuations in ^{14}C within the Lake Suigetsu record (and even more so in sparse ^{14}C from marine sediment cores) as denoting definitive atmospheric ^{14}C-age plateaus. Note that, even with the Lake Suigetsu core on its own, we cannot confidently identify plateaus using the IntCal Bayesian spline approach. However, we wish to make explicit that we are not intending to exclude the possibility that atmospheric ^{14}C structure does exist in the period from 55 – 14 kyr cal BP. But we also stress that identification of this potential atmospheric structure is not sufficient for PT.

It is quite possible that, as we get more precise atmospheric ^{14}C data from 55 – 14 kyr cal BP, we can identify more genuine atmospheric ^{14}C structure, notably with additional records based on tree-rings (e.g. Cooper et al., 2021). Indeed, we hope for this to be the case and are working towards this goal. However, even with reliable atmospheric ^{14}C structures identified, this would not get around the larger and more intractable issue of identifying, matching and tuning any such genuine atmospheric structure within marine sediment cores and extracting at the same time, a priori unknown MRA variations.

We wish to stress again that our main argument is not about whether, or not, there are periods where the atmospheric ^{14}C-age has structure and variation; or that the radiocarbon clock may tend to move slightly faster/slower in periods. Further, we agree with the general concept of wiggle matching – although disagree as to the equivalence with PT.

Our main argument against PT relates to the detection and absolute dating of equivalent ^{14}C structures in marine records – data without an accurate timescale, characterized by sparse ^{14}C

data affected by an unknown MRA, which shifts and smoothes the ^{14}C structures by unknown amounts.

Page 14 (SG): Comparative tests (Sarnthein et al., 2015; Fig. 2a) revealed good agreement of plateau boundaries deduced by visual inspection and by calculating the 1st derivative, though differential analytical errors have not been considered in these tests.

We thank the authors for this clarification that the different ^{14}C analytical errors are not taken into account in the 1st derivative technique introduced by Sarnthein et al. (2015) as it was not clearly stated in the original publication. We note that analytical error bars should be taken into account because a kink or a plateau in a ^{14}C vs. depth graph could also be a structure within the analytical errors and PT would be useless in this case.

About our statistical simulations of plateau tuning

Page 17(GS): The choice of a gradient of 1, that is one ^{14}C year per calendar year instead of 0.5 seems better as it approaches natural ^{14}C decay and the IntCal curve for this time interval. However, this choice should not change the results; just make them more difficult to read in the figures.

Page 18 (GS): As B&H state in the discussion of this experiment, they tried to mimic the plateau tuning of the papers by Sarnthein et al. Yet, the details of the kernel set-up determine the smoothing and their kernel optimization may have been slightly different.
... On the contrary, Fig. 6 shows that even in a short segment with only three plateaus two out of three features of the underlying fine structure can be recovered and correlated

Following the comment from RC1 (Paula Reimer) we have updated our statistical simulations to test the influence of the gradient threshold (0, 0.5 or 1) and of the kernel bandwidth (see above the new Figures and associated text in our responses to RC1). We demonstrate that both parameters have a strong influence on the PT results, leading to non-robust results for the chronology and MRA reconstruction.

We also stress that a gradient of 1 is not a natural choice for defining a plateau. Indeed, a gradient of 1 is what the slope (^{14}C yr /cal yr) should be without any perturbation of the radioactive decay.

One must also remember, that our marine sediment simulation study only looked at the effect of a lack of initial calendar scale on which to estimate ^{14}C-age/cal-age plateaus and sample sparsity. It did not consider the effect of variable MRA which will add highly significant confounding to the identification of ^{14}C-age plateaus in the marine sediment cores.

Miscellaneous

Page 7 (SG): We welcome B&Hs' notice of numerous, though unlikely abrupt rises in atmospheric ^{14}C. Also, we ourselves have discussed them internally already over years and regard them as valuable novel signals of short-term variations in ocean-atmospheric carbon exchange. As B&H say, processes controlling these partly fairly instantaneous processes are complex and difficult to model but highly challenging and worth to be traced by future studies.

We thank the authors for acknowledging our contribution about evidencing the implication of numerous instantaneous rises of atmospheric $\Delta^{14}C$. We are happy that they agree with us that these events are unlikely and difficult to model.

In their numerous papers since 2007, the authors focused on the ^{14}C age plateaus, only representing them in terms of ^{14}C ages versus calendar age or versus depth, but never in the $\Delta^{14}C$ space. This is what we did in Fig. 3 of our original submission, which led to the puzzling implication that atmospheric $\Delta^{14}C$ must have jumped abruptly and frequently (if the ^{14}C behaved according to the underlying hypothesis of PT). These instantaneous $\Delta^{14}C$ rises ranging from 50 to 250 ‰, correspond to the very short periods between the ^{14}C age plateaus.

It is indeed these inter-plateaus events that pose the biggest challenge in terms of their implication on the underlying physical mechanisms. There is no known mechanism that could be responsible for such abrupt and large asymmetric wiggles of the atmospheric $\Delta^{14}C$. Instantaneous ^{14}C production increases, that result in about 4 times the average production in a year, were discovered recently by Miyake et al. (2012), but the size of these spikes attributed to extreme solar particle events is an order of magnitude smaller in terms of $\Delta^{14}C$ than that required to explain the jumps implied by PT. Furthermore, there is no evidence of huge corresponding spikes in the ice core ^{10}Be record. As underlined above in the section about box-modelling, it is also unlikely that changes of the carbon cycle are responsible for such large, frequent and instantaneous $\Delta^{14}C$ rises.

Page 13 (SG): Consequently, we do not see any need either "to smooth Suigetsu-based records" and/or to adjust the marine MRA to a marine average MRA of 480 yr used by the authors.

It is correct that the nature of statistical regression techniques (such as the Bayesian spline used to create IntCal and also the Suigetsu-only curve shown in Figs 1-3 of our original submission) are to somewhat smooth the underlying raw ^{14}C observations. However, it is well recognised that such smoothing is needed to reduce spurious variance in the resultant curve estimate. The underlying ^{14}C measurements are observed subject to noise which will typically mean they are more highly scattered than the underlying atmosphere. Should no smoothing take place, the curve will tend to overfit the data and introduce features which are simply artefacts of the observational noise as opposed to genuine atmospheric signal. Such noise artefacts should not be used in determining an atmospheric tuning target. We therefore maintain an argument some smoothing is required when using ^{14}C determinations to reconstruct atmospheric levels – our Bayesian spline aims to provide a level of smoothing which is adaptive to the data it is fitting.

In our submission, we did not adjust the MRA in our simulated marine sediment cores to a value of 480 ^{14}C yr but to 400 ^{14}C yrs, so are somewhat confused as to this comment. Further, this constant MRA adjustment has the effect of simplifying the identification of ^{14}C-age plateaus compared to the variable MRAs inferred by Sarnthein et al.

Note that the 480 yr value corresponds to the dead carbon fraction (DCF) used for the speleothem record from Hulu Cave in China. In a way similar to our box-model, the carbon transport and mixing processes, passing through several soil boxes, should have somewhat smoothed and filtered the atmospheric ^{14}C variations recorded in the speleothems.

Work on how to best model this speleothem filtering is ongoing and will hopefully lead to improvements in the next calibration curve. We also hope for more directly-atmospheric ^{14}C records in the future (such a floating tree-ring sequences and macrofossils similar to Lake Suigetsu).

Page 9 (GS): It would have been great to have a scientific in-depth discussion earlier, preferably in the discussion phase of the synthesis paper, or earlier.

We would like to note we submitted our paper on Jan 31st of 2020, within the open discussion phase of the paper by Sarnthein et al. (2020). For reasons outside of our control this unfortunately did not make it online at that time.

We note also that intense discussion and debate about PT has also taken place during meetings of the IntCal working group and focus group on marine archives (Belfast 2010, Paris 2012, 2015, Utrecht 2016 during which Dr Grootes proposed and explained the technique).

Additional references (all the others are already listed in our submission)

Ausin, B., Sarnthein M., Haghipour N., Glacial-to-deglacial reservoir and ventilation ages on the southwest Iberian continental margin. *Quaternary Science Reviews,* 255, 106818, 1-12, 2021.

Cooper, A., Turney, C.S.M., Palmer, J., Hogg, A., McGlone, M., Wilmshurst, J., *et al.*, A global environmental crisis 42,000 years ago. *Science* 371 (6531), 811-818. 2021.

Corrick, E.C., Drysdale, R.N., Hellstrom, J.C., Capron, E., Rasmussen, S.O., Zhang, X., *et al.*, Synchronous timing of abrupt climate changes during the last glacial period. *Science*. 369, 963–969, 2020.

Siani, G., Michel, E., De Pol-Holz, R., DeVries, T., Lamy, F., Carel, M., Isguder, G., Dewilde, F., Lourantou, A. Carbon isotope records reveal precise timing of enhanced Southern Ocean upwelling during the last deglaciation. *Nature Communications* 4, 2758, 2013.

Skinner, L. C., Waelbroeck, C., Scrivner, A., & Fallon, S. Radiocarbon evidence for alternating northern and southern sources of ventilation of the deep Atlantic carbon pool during the last deglaciation. *Proceedings of the National Academy of Sciences*, 111, 5480–5484, 2014.

Skinner, L.C., Freeman, E., Hodell, D., Waelbroeck, C., Vazquez Riveiros, N., & Scrivner, A. E. Atlantic Ocean ventilation changes across the last deglaciation and their carbon cycle implications. *Paleoceanography and Paleoclimatology*, 36, e2020PA004074, 2021.

Weninger, B. and Edinborough, K. Bayesian ^{14}C-rationality, Heisenberg Uncertainty, and Fourier Transform. *Documenta Praehistorica* 47:536-559, 2020.