Response to Referee 1 – Brummer et al., "*Modal shift in North Atlantic seasonality during the last deglaciation*"

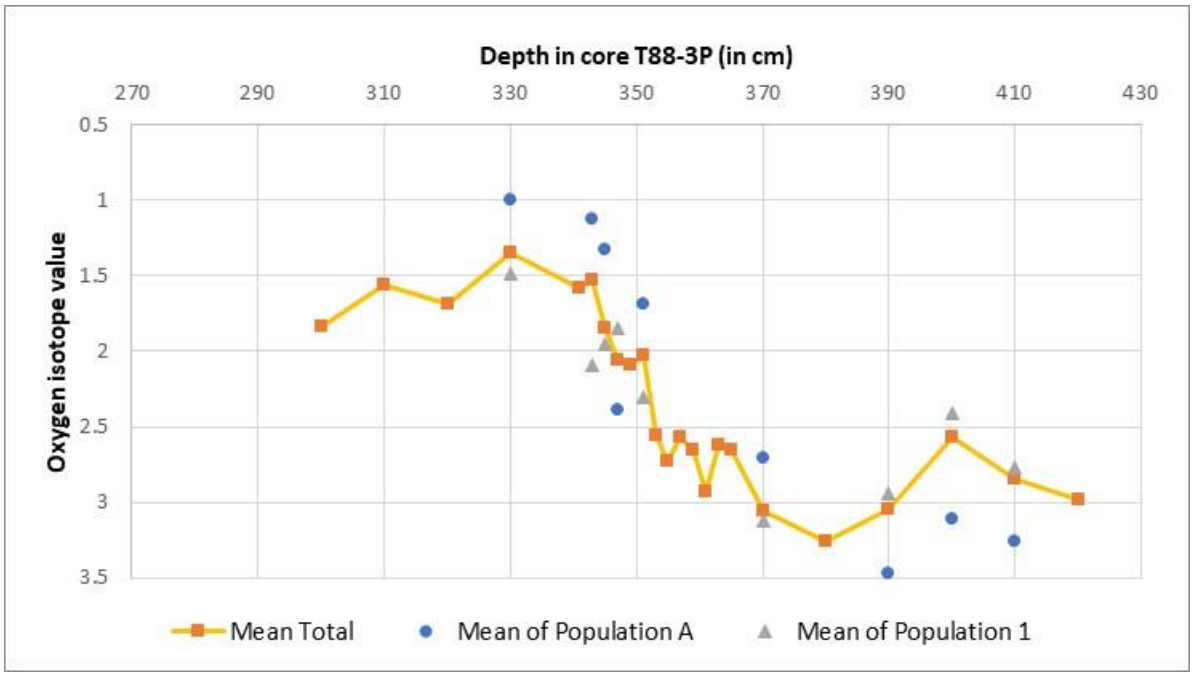
We thank the reviewer for raising some important points. As the reviewer notes our conclusions are not altered by our age model however upon reflection we agree with the reviewer and therefore would like to take the opportunity to expand our manuscript's age-depth model (including adding SAR) in a revised Manuscript (we thank the reviewer for drawing attention to the oddity in table 1 and will correct this). We also thank the reviewer for noting that our $\delta^{13}C$ of *G. bulloides* is actually normalised between 0 and 1. The 'normalisation' has kept the difference between the absolute values, the data is just presented on a relative scale. We will correct this is in a revised MS (although we present the absolute values in a plot below). In the following, reviewer comments are in RED, our responses are in BLACK:

There are three major concerns that I have and which I will outline first.

1) Unimodal mode of G. bulloides and G. bulloides δ13C values

The authors state that the single specimen isotope data of G. bulloides are unimodal, but give not reasoning for this statement. Subsequently, they use the unimodal distribution of G. bulloides as evidence that the two populations of N. pachyderma cannot be related to bioturbation (more on this in point 2). I would like to see some justification for declaring the G. bulloides data unimodal in the text. Whereas the δ18O values show much less scatter than the N. pachyderma data, the respective δ13C data show a range of 0.5‰ at some levels and I wonder, if this is not a reflection of more than one population.

We did not produce a similar figure of the unimodal nature of *G. bulloides*, however at the reviewer's suggestion we have performed this and there are at some depths indeed more than one population – this is an equally interesting result and we will add this to a revised manuscript. We thank the reviewer for their suggestion – our focus was on *N. pachyderma* for this analysis because as a polar species it should have a reduced ecological range and hence our interest in more than one population. A quick figure of the *G. bulloides* data is plotted here:

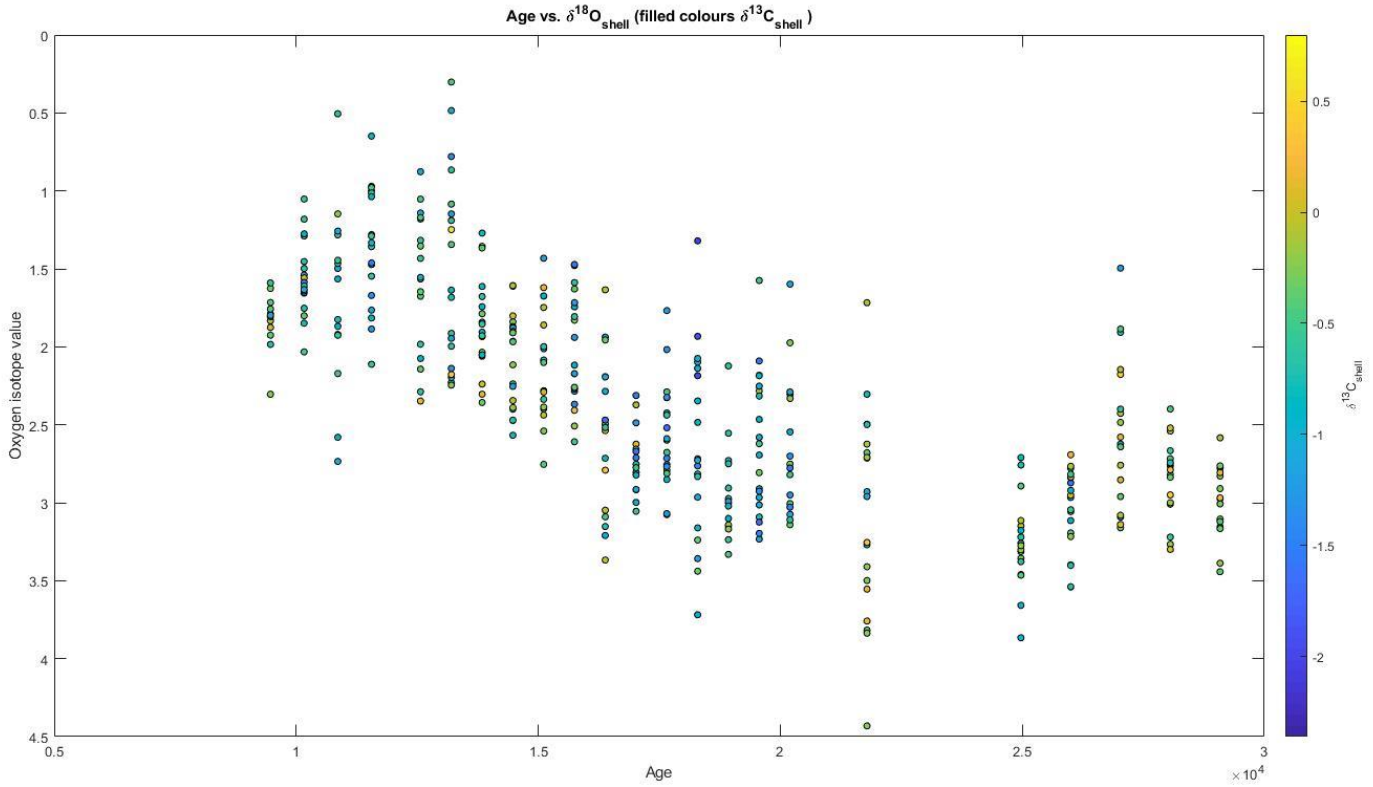


In addition, we will add in values of glacial *N. pachyderma* and *G. bulloides* from much deeper in the core which we can add as a comparison between ecological change across the deglaciation and a glacial interval.

It is important to clarify that: "Subsequently, they use the unimodal distribution of G. bulloides as evidence that the two populations of N. pachyderma cannot be related to bioturbation" our exclusion of bioturbation is not only based upon our perception of the unimodality of *G. bulloides* but as we state further down in the same section: "*However, we exclude this particular scenario because sedimentary features (Fig. 2) indicate a lack of discernible mixing, i.e. the sharpness of the IRD percentage, the Log(Ca/Ti) and the percentage of NPS all indicate that bioturbation is at a minimum*". And hence (This statement is, however, only valid) is not the sole reason for whether or not our statement is valid, although it is a strong argument.

This statement is, however, only valid if the δ13C values plotted in Figure 3 are actually correct, because G. bulloides δ13C values should (mostly) be negative and the scale on the Figure is positive and has exactly the same range as for N. pachyderma.

Apologies, we thank the reviewer for pointing out this mistake. Whilst, the values were correct, the plotting tool had rescaled the colour scale to values between 0 and 1. We will correct this in a revised version, but for now we present the data not scaled.



2) Influence of bioturbation

Whereas I agree with the authors in the general sense that the occurrence of two populations cannot be explained by bioturbation, I would urge them to be more careful in those cases where one of the populations is presented by only 1 to 4 specimens.

We agree, hence why we sought to give the reader alternative explanations as well, i.e., section 4.2 onwards. There are only 3 populations with less than < 3 specimens; only 5 populations with less than < 4 specimens. We will add in the following text: "However, it is important to note that for several depths in core this second population may only represent a few specimens ($n_{< 3 \text{ specimens}} = 3$; and $n_{< 4 \text{ specimens}} = 5$)"

In this regard, it is essential to include an abundance record (which could be the N. pachyderma ratio record from Fig. 2) of both species in Figure 3. Since Figure 2 is presented vs. depth and Figure 3 vs. age, it is impossible for the reader to see where abundance minima of the respective species could have led to a "bias" in the single specimen isotope data (also in G. bulloides during periods of near dominance of N. pachyderma).

Although we did present both (top panel) depth and (bottom panel) age in figure 4, we can certainly add additional panels into the figures to highlight the abundance of the species. We will plot the abundance data also on both age and depth scales.

For example, I do not perceive the argument of the unimodal mode of G. bulloides valid for the two specimens of population 2 in the third line of Table 2 [see note below on correcting column 1 of this table], if that level has already a low abundance of N. pachyderma and can thus be much more likely affected by –even if assumed minor, i.e. over 5 instead of 10 or 20 cm depth– bioturbation.

We thank the reviewer for their comment, though this is why we state, "alternative scenarios that give the same or a similar solution for the existence of two populations can be envisaged". Whilst we have explained (section 4.2.3) how bioturbation would potentially affect our observations down core – we can present the abundance data that we do have and include a discussion of the abundance of foraminifera with respect to bioturbation.

In addition, Figure 3 should include a plot showing the variations in the sediment rates, so that the reader can see where low sedimentation rates might have increased the chance of bioturbational mixing. Including these plots might not change the story, but provides the reader with the option to judge him/herself in which levels bioturbation might have affected the single specimen data (and to what degree) or not.

We have a sparse number of tie-points, which is why we did not plot the sedimentation rate in one or more panels (in the background). A sparse number of tie points may give a spurious impression, for instance at times of high or low IRD the SAR may vary considerably yet with a sparse number of tie points we have only 'book-ended' these results with a single SAR value. Figure 4, in which one panel has age and the other depth was our attempt around presenting the depth to the reader (as the reviewer suggests). Here the reader can see the two populations vs depth in core and therefore can for themselves consider the mixed layer or bioturbation depth which may or may not vary between 5-15 cm. With the reviewer's suggestion of expanding figure 4 (see comment above) we hope that those changes will be sufficient.

Additional comments:

Main manuscript p. 3 abundance counts: please specify a) how the % IRD was calculated; b) why a Ratio of NPS was calculated and not the more commonly used % N. pachyderma.

IRD was calculated from a sum total of foraminifera and IRD. This does have complications for the calculation of % N. pachyderma as it is a closed sum with some variation due to changes in IRD. Whilst this is less than ideal, unpublished data comparing these methodologies shows that the % N. pachyderma produced from a sum of foraminifera and IRD is consistent with %N. pachyderma as a sum of only foraminifera, when IRD is less than ~50% of the total grains. Higher values of IRD will, of course, alter this.

p. 3 Stable isotope section: please mention a) the resolution at which the single specimen measurements were done (4 cm?); b) if the N. pachyderma specimens were encrusted; c) which are the international carbonate standards used during the stable isotope analyses?

We performed faunal counts every 4 cm (line 2, pg 3). We will clarify that this spacing is different for the isotopes. Therefore, we intend to alter pg. 3, line 2 as follows:

"The core sections of the entire working half were sampled every cm, resulting in 1 cm sample slices that were each washed over a 63 µm sieve mesh, dried overnight at ~75°C and subsequently size fractionated into 63-150 µm and >150 µm. For abundance counts of planktonic foraminifera, slices every 4 cm were used, the counts were performed on…"

We will alter pg. 3 line 13 to: "Slices for isotope analysis were selected first at 10 cm resolution and then at specific sections down core every 2 cm. For each slice 20 shells of both left coiling *N. pachyderma* and *G. bulloides* were picked at random from the 250 - 300 µm size fraction (Figs. 2-4)."

p. 3 core stratigraphy (besides comments above on 14C calibration): may be specify that you follow Reimer et al. (2013) when using ΔR of 0±200 yr.

We will repeat the reference at the end of the sentence, so that sentence will read: ", using the Marine13 Calibration curve (Reimer et al., 2013) and a reservoir age of 400 14C years with an error of 200 14C years, expressed mathematically as ΔR: 0 ± 200 14C yr (Reimer et al., 2013)."

line 29-30: if you keep the sentence, specify which sample was excluded (do not assume that every reader will read the supplementary material in detail).

We will make a note in the table to show which is excluded.

line 31-32: how many specimens of G. bulloides and G. glutinata were analyzed for the "bulk" analyses?

The data is based upon the mean of 2 groups (comprised of 5-10 specimens per group) for each species – we will reiterate this in the paper.

line 35: include that the tuning was done to the δ18O record of NGRIP, which, I assume, is presented on the GICC05 chronology. If you used NGRIP on GICC05, did you remember to correct the GICC05 b2k ages to BP ages (by subtracting 50 years) to make the tuned ages compatible with the calibrated 14C ages?

We will adjust the figures accordingly, as stated in the supplement we use an earlier chronology – we will therefore replot and alter the age model according to the GICC05 chronology (this shifts the age ever so slightly).

line 36-37: you are providing information on temporal resolution and not sedimentation rates. I do not find this very informative and would like to see a figure showing the variations. Also, the sentence in its current phrasing is incomplete.

We will add a figure and complete the sentence.

p. 4 line 4: what does IFA stand for?

IFA stands for Individual foraminiferal analysis – we shall, alter the header to: "1.1. Seasonality and single foraminiferal analysis (SFA)" and then alter the header of 2.5 to the same acronym.

p. 4 line 20: year missing for Jonkers and Kucera reference

We will add the date

p. 5 line 14-15: what about within glacial mixing/bioturbation?

The detection of bioturbation is intrinsically related to the difference between two samples, if two samples with uniform values between them are mixed, then it would be impossible to distinguish them, though it does not mean it does not exist. We will clarify this

p. 6 line 35: N. pachyderma δ18O data not shown in Figure 2.

Indeed. Consequently we will change "…of either N. pachyderma or G. bulloides (Fig. 2)…" into "…of either G. bulloides (Fig. 2) or N. pachyderma (Fig. 4)…"

Table 1: following the recommendations of Stuiver & Reimer " Users are advised to round results to the nearest 10 yr for samples with standard deviation in the radiocarbon age greater than 50 yr".

We will round these numbers for the main text and add a supplementary table of unrounded numbers (whilst we agree with the referee we also wish to allow for future readers to know the actual number used).

Table 2: first column: please correct; what you are listing are not or incomplete depths. since the data itself is not shown vs. depth, it would be good to have an age column as well. Reduce the number of decimal places in the Prob and Mean columns, so that the numbers become easier to read.

We will alter this accordingly (it was the sample ID). We will also round the mean, standard deviation and prob numbers to 2 decimal places.

Figure 3, 4, S1 etc.: in all the axis label referring to the NGRIP δ18O data, replace the "SW (sea water ??)" by "ice". Provide reference for NGRIP data in figure captions.

We will alter to just $\delta^{18}O$ as VSMOW is already indicative of the substance.

Figure 3: as mentioned already above under point 1, correct the δ13C scale for G. bulloides.

Altered accordingly.

Inconsistency between p. 3 line 30, supplementary material: you state that the deepest/oldest 14C age was not used/excluded; so why it is then shown and used in Figure S2

We will clarify in a revised MS; we exclude it not for it being incorrect or wrong but due to the fact that the calibration curve at the older end is based upon 'noisy' data (this is not critique, rather it is 'the best of a bad lot' and just a comment on the underlying data used in the construction of the calibration curve) therefore whilst its exclusion as a tie-point in an age model is circumspect it can still be used as an indicator that the age model appears to be 'working' (it is not blank, for instance, therefore the pooled age is likely not older than 50,000 years).

As outlined above we agree with the reviewer that we could be more clearer with our age model, and that this section could benefit from being moved into the main text. Therefore, we will make the following changes that the reviewer in these comments has suggested:

3) Age model and 14C calibration

The authors made the effort to test different approaches to establish an age model, but in the end the reader does not know, which age model/age control points were used to produce the record of the data vs. age as shown in Figure 3. So please, specify this and provide either in the main manuscript or in the supplementary material a table listing the final age control points. Did you combine? If yes, did you then discard some calibrated ages?

We agree with the reviewer that we should expand our age-depth model – in a revised MS we will add a section of the text to include a more detailed discussion and explanation of the age model.

Issues with the text and information in Table 1 regarding the 14C calibration: Table 1 and section 2.4 and supplementary material: your measured age should be the same as the conventional age, i.e. the raw 14C concentration converted into an uncorrected 14C age (using the Libby half-life). If you calibrate with Marine13 this uncorrected age would be the one used to calibrate. So I do not understand how your Table 1 can list conventional ages that are 400 years higher than the measured age –which to me looks like a reservoir age correction going into the wrong direction! And I am not sure, which age –measured or conventional– was actually calibrated! If you analyze marine material like foraminifera the measured/conventional age needs to be corrected for the reservoir effect, i.e. transferred to "atmospheric 14C levels" by subtracting the reservoir age (such as 400 yr), if you want to calibrate with atmospheric level calibration data like Intcal13. Since you are calibrating with Marine13 you do not use a fixed reservoir age (of 400 years)! During the Holocene (0-10.5 cal ka BP) section the reservoir age is provided as outcome of the ocean-atmosphere box diffusion model and varies "significantly" over time –see for example Figure 4b in Hughen et al. 2004 on Marine04. In the glacial section, where a fixed reservoir age is used, the value is 405 years and not 400 years (see p. 1877 in Reimer et al. 2013). Inconsistency between p. 3 line 30, supplementary material: you state that the deepest/oldest 14C age was not used/excluded; so why it is then shown and used in Figure S2? While correcting the 14C calibration will change the age model, this will not affect the general conclusions of the manuscript.

We thank the reviewer for noticing the discrepancy – and will alter the text accordingly.

Supplementary material text: line 24 insert δ18O before ice core and mention that the NGRIP record is on the GICC05 time scale.

The data is based upon NGRIP (North Greenland Ice Core Project members et al., 2004), in a revised MS we will adjust this to GICC05.

Figure S4: the right panel does not show the filtered NGRIP record = tuning target. Why is the SPECMAP error applied and not the GICC05 errors?

We used a perceived tuning error which SPECMAP calculated – not the NGRIP error – as this would give us how the signal migrates (i.e., atmospheric signal vs. ocean signal). The error for a core will naturally be larger than the NGRIP error, as one is a slow 'responder' whereas the other is a fast 'responder'.

line 27: provide more information on the "simple filter". for which frequencies did you filter and why?

We used a filtering algorithm (~500 year time window), producing a series of filtered variants (max.; min.; mean; and so forth) of the time series to reduce the variability from a high-resolution time series (NGRIP) to one that shows the major long-term changes. This may appear counter intuitive; however, this is (i) to reduce the effect of over tuning of small-scale high frequency variation and (ii) to produce an NGRIP signal that would be similar to a down core record (i.e. a smoothed signal from a high resolution signal).