

EGUsphere – May 2024

Western Indian Ocean bottom water temperature calibration – are benthic foraminifera Mg/Ca ratios a reliable paleothermometry proxy?

Larsson, V. & Jung, S.

Overall Review:

This study presents some new Mg/Ca (and other elemental/Ca data) from benthic foram samples from core-top locations in the Western Indian Ocean basin and compares it to previously reported values from other locations in the wider Indian Ocean and beyond. They walk through three experimental methods for cleaning prior to Mg/Ca analysis and compare results in the context of contamination and data exclusionary procedures. They show agreement with some previously published calibrations for Mg/Ca – Bottom Water Temperature. This is a significant amount of work and the authors do have some exciting things to share.

However, their data is quite limited after removal of samples with apparent contamination, such that their final calibration model only includes 4 samples. As far as a method-based paper discussing the range of cleaning methods and implications of those methods on data collection, this paper is valuable. As far as a new model for reconstructing temperatures, this data should be taken with caution due to such low final sample numbers. However, their data do seem to fall within previously published values which shows good continuity. They could stress this more in the abstract as well, and make it very clear that their new calibration model contains only $n = 4$ samples.

I believe the paper should be published, but first with edits. Specifically, some of the text is confusing to follow as written, with many subsections in the discussion that are quite disparate/not well connected. If the authors are able, it would be good to go back through the discussion in particular to draw out and make very clear the major findings of their work. In general, a clearer outline of what this study brings to the literature/scientific field and the suggestions/findings for future work is needed. The Summary and Conclusions section is quite sparse in this regard and could be used as a place to discuss their findings in greater depth.

With respect to the methodological testing (cleaning methods), it would be helpful if the authors can work to make it far clearer what the difference between experiment 1 cleaning versus experiment 2 versus experiment 3 cleaning methods were – it's difficult the way it is worded now to follow. Perhaps a table outlining the differences of each procedure, and which steps were included in each so people can easily cross compare the methods would be instructive. Is the only difference that the methanol washes timing is reduced from 1-2 min to 20 seconds? Is it expected that the cleaning procedure would work the exact same way on the benthic species? Can the authors make a note of their expectations on this and reasoning for using *G. ruber* as their cleaning methodology species?

Please also update figures throughout to make it clearer for the reader to distinguish between datapoints, as they stand now many of the symbols are too small and so similar to one another.

Overall, it is difficult to determine whether they conclude that benthic foraminifera Mg/Ca ratios are a reliable paleo thermometer in the Indian Ocean. Please be more explicit in your conclusions about your findings and how they fit within the larger context of benthic foram Mg/Ca paleothermometry.

Detailed Review:

Figure edits:

Figure 1 – maybe label with (a) and (b) so that it's clear beyond colour which map is showing which core-tops.

Figure 2 – label with (a) and (b) (c) so it's clear in your figure caption when you are speaking about temperature and salinity and the transect map.

Figure 3 – Be very clear in your figure caption. Indicate when you are referring to the 'blue arrow' and 'red arrow' because readers may think you are referring to the water temperature blue and red colours.

Figure 4 – Perhaps it might be easier to tell the difference between crushed versus not crushed samples if you colour them differently or have one filled in one empty triangles?
Also - does it really make sense to have one regression line that includes both crushed and un-crushed data points - what about two separate regressions showing the relationships across treatments?
Figure 4 caption: Is there a way to quickly describe the procedure here instead of referring to it as from Experiment 2? perhaps you can say using the method with shorter methanol cleaning step?

Figure 6 - In all species evaluated across all cores? indicate this if that's the case. Otherwise, where (which cores) are these data from?

Also – panel D doesn't exist – where you indicate in your caption there should be the correlation between total Ca and contamination (Fe/Ca)

Rather than just saying 'horizontal lines' say "Orange lines show..." and indicate which value corresponds to which Element/Ca ratio.

Figure 8 – The legend box is outside of the figures, try to align so it does not cross the figure border the way it does now.

Figure 9 – How many samples make up each relationship – indicate sample number with " n = "
Capitalize the C. for *C. wuellerstorfi* on figure

On line 329: should also include (u) here after Mg/Ca for (uncontaminated) indicator you have in your figure legend.

Figure 10 - So all calibration models/ data presented are from this one species? make sure that is clear in the text above

Include reference/citation for the S. W Indian Ocean grey samples in your figure legend as you do for the purple and blue.

Line by line edits:

- 14:** the word entailed doesn't seem appropriate. Perhaps the word elucidated?
- 19:** with what error can the 'wider' Indian ocean calibrations be used in the Western Indian Ocean? Could report that error here for brevity and maximal impact of abstract
- 22:** With what error can BTW be reconstructed? What does your calibration error translate to in degrees C error of the BTW?
- 24:** remove the word 'the' after controls...
- 25:** the re-distribution of heat **in the oceans** is an
- 28:** sentence should read: "... of the sensitivity and changes in thermohaline circulation. For example, on glacial-interglacial time scales, NADW and AABW..."
- 34:** can use $\delta^{18}\text{O}$ here instead of 'stable oxygen isotopes because you've already defined it earlier
- 41:** perhaps instead of "being developed" saying "is still unresolved" is better?
- 43:** Before starting in on the discussion of Mg incorporation, make clear that forams make calcite tests which is a calcium carbonate matrix... and describe that Mg substitutes for Ca in the lattice.
- 43:** remove the word "also" after "Mg²⁺ are also incorporated..."
- 44:** indicate that Mg/Ca does not just depend on Mg/Ca of seawater and elemental partitioning...it's an endothermic process that relies on water temperature – hence the reason we can use it to reconstruct water temperatures!
- 48:** Define BWT here before using the acronym as it has not yet been defined in the text.
- 49:** Temperature appears to be the dominant environmental factor driving what? Indicate you mean the incorporation of Mg?
- 52:** carbonate ion saturation being dominant of/at what? What process is it affecting? Mg incorporation? explain
- 54:** What do you mean by 'spatially-varying' please elaborate here.
- 71:** "larger lowering" is awkward wording, consider rephrasing this
- 80:** "... determining **bulk or whole specimen** calcite Mg/Ca ratios"
- 84:** "*Cibicidoides wuellerstorfi* **has** been one of..."
- 85:** provide example citations for the use of benthic species for stable $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ reconstructions.
- 87:** when referring to Mg/Ca incorporation – it is better to say Mg/Ca signatures OR Mg incorporation.
- 91:** "... being usable..." is an awkward way of saying this. Perhaps "employed"
- 92:** "Uncertainties remain (Stirpe et al. 2021)" – be clearer here, what uncertainties are you referring to that are raised in Stirpe et al. 2021.
- 92:** "... entailing the need..." entailing is an awkward word choice: perhaps 'pointing to the need' or 'elucidating the need'
- 94:** remove the word help before 'improving' and change 'improving' to 'improve'
- 95/96:** you compare to calibrations from the Indian Ocean – are these previously published calibrations? If so list the citations if they are not too lengthy?
- 108:** should say... "...core-top transect is comprised of..."
- 108:** you only indicate 2 water masses in this first list... and go on to introduce the third later but it is slightly confusing/misleading to say three here and not list all three
- 113:** where are the CTD temperature data from? When were they collected relative to the core-top collections?
- 134:** no need for brackets around "where possible" when you are using commas
- 144:** NIOP929 – what is this? Which core? From what cruise/expedition/researchers?
- 145:** The sentence should read: "The samples were wet sieved over..."

- 147:** The sentence should read: "...remove silicates, a hydrogen peroxide treatment to remove organic matter, and followed by an acid ..."
- 149:** "depend" should be "depends"
- 153:** The sentence should read: "...except for a reduction in duration of the methanol washes (25 seconds...)"
- 154:** The language "following previously analysed samples in the laboratory." Is not clear – please clarify and/or rephrase.
- 159:** for how long was the hydrogen peroxide treatment? At what temperature?
- 166:** Does this also mean you lose Mg in the glass slide method as well as Ca in order to keep the ratios similar to the pin method this must be true?
- 178:** If the data in Figure 4 below show experiment 2 results, then pointing to it here for experiment 1 calcium values doesn't make sense
- 183:** Did you run a sensitivity test/ leverage test to see how much leverage that outlier has on your regression/model? if it's low then you can likely include it, if it's high then it's likely skewing the results.
- 184:** Is it really true that there is no relationship between Al/Ca and Mg/Ca? or is it just non-linear like logarithmic? Looks like it would be.
- 185:** The correlation for not crushed is only really strong if the outlier is included, otherwise just clusters... the correlation for crushed tests is linear however
- 224:** Please describe in more detail what the standards were - you should have that information from the lab in which these samples were run and it's standard practice to include the type/name of the specific standards used.
- 225:** You should rename this Figure B1 - to indicate you are referring to Figure 1 of Appendix B and not Figure 1 of the manuscript.
- 229:** 'effect' should be 'effects'
- Section headings 3.1 and 3.2** - you refer to more than just Mg/Ca ratios in these sections, perhaps re-naming to say elemental ratios or something similar?
- 261-264:** It does not look like the samples are below the contamination thresholds? you even indicate in table 2 which samples are eliminated due to contamination.... so this statement can't be true.
- 298:** r^2 is a different metric from R^2 – you have used both throughout the text, so go through and ensure
- 310:** Perhaps a sentence following indicating how many samples remained for the core-top calibration is useful here.
- 318:** capitalize the 'f' in figure 9 – Figure 9
- 323:** This calibration is based on 4 samples, one of which is quite far away from the others - could be contributing significant leverage to the calibration... i.e. without it the entire relationship would fall away. Have you tested this? Also, if you are going to report this calibration model (or any throughout the manuscript) it is good practice to indicate the number of samples/data points you have that built the model.. in this case: $n = 4$.
- 348:** Maybe rather than 'abnormally' say anomalously high, not abnormally since you have cited evidence where similar reports have been made
- 350-354:** So did you exclude this sample from calibration models? make it very clear if so.
- 355- 360:** So did you exclude this sample from calibration models? make it very clear if so.
- 362:** Are these studies also reporting only c. wuellerstorfi calibration models or which species are they reporting? indicate this here if they are species-specific or mixed-species models!

369: you have ??? after Figure in your reference to figure 11

393: the word maybe should be two words in this context: may be

426:) close (Figure 6) with a bracket

428: under what subsection/subtitle number are you pointing readers to for discussion on degree of contamination?

474: This first sentence is awkward wording...do you mean to say your core top samples are closely located to previously published data?

514: risk should be risks

517-518: How? if you are to suggest more work should be done it's great if you can point to examples of next steps

520: benthic/planktic should be benthic versus planktic

526-527: Also include that this calibration only includes 4 samples.

534-535: It would be great to say something more about your study.... while you are not able to speak on the other uncertainties/limiting factors, you have shown here that cleaning procedures may need to be refined and conducted species-by-species.