RC3

In “Western Indian Ocean bottom water temperature calibration – are benthic foraminifera Mg/Ca ratios a reliable palaeothermometry proxy?” Larsson & Jung measure Mg/Ca ratios in benthic foraminifers across a depth / space transect in the Western Indian Ocean to determine if a local Mg/Ca-temperature calibration is appropriate. They also test several cleaning techniques to determine their effects on contamination, sample yield, etc.

I’m no expert in Mg/Ca paleothermometry, but it seems like the half of the paper devoted to cleaning procedures is thorough and useful (if a little hard to follow). The calibration, however, is based on very few points (and anchored by one high-temperature sample), which the authors acknowledge, and is not likely to be used on its own. I think the exercise is still valuable, given that most of their data fall in the Mg/Ca-temperature space in prior calibrations, but the manuscript needs a clearer through-line, e.g., “best” cleaning protocol established -> despite efforts, majority of the specimens contaminated -> resulting calibration is sparse, but (most of) the data seem reasonable and species- / habitat-specific contamination thresholds, calibrations, etc. are recommended.

I believe that this can be of use to the paleoceanographic community pending major revisions in terms of structure, organization, and clarity – I would also highly recommend a thorough grammatical overhaul, there are numerous minor issues only some of which I’ve noted below. I don’t believe further laboratory analyses are required, although there are one or two instances where they might be helpful.

Thank you for your valuable input and for taking the time to thoroughly review our paper.

Your input on organisation and structure will be taken into account. We have aimed at making a balance between the methodological part of the paper and the results that as you have mention might still be of use despite contamination as some of the results agree with previous studies. However, with regard to the cleaning methods, we have not tested a wide range of approaches. We have followed one established cleaning procedure and adjusted the timings to accommodate for potentially more/less contamination in samples.

General comments:

The title is rather vague and doesn’t reveal much about the study’s true findings. Maybe something like “Persistent contamination issues preclude a simple benthic Mg/Ca-temperature calibration in the Western Indian Ocean?” I’m sure you can come up with something better.

Thanks for this useful suggestion. The next version will have an improved title.

Your tests of the various cleaning procedure parameters are a major part of the paper, and I would mention it in the title. I think you need to emphasize that this is a valuable contribution to Mg/Ca thermometry, however – it reads to me like a sidenote compared to the calibration until you reach the later part of the manuscript.

If I missed this, I apologize, but any ideas as to why the other Indian Ocean calibrations didn’t have as extensive of contamination issues?

The tests of the various cleaning procedure parameters are indeed a major part of the paper however, as the methodology that is used is closely following the methodology developed by Barker et al. 2003, the main point of the paper is to highlight how the methodology by Barker et al. 2003 is adopted to

different samples and that more research is needed to explore what adaptations are needed for different samples, it also proves that calibrations comparison might be more difficult if different adaptations have been made in response to varying contamination.

Since no comments on adaptations to Barker et al's (2003) methodology have been mentioned we have assumed it has followed exactly the procedure of Barker et al., 2003. We don't know why previous papers have not run into these issues of contamination, as the cleaning methodology , especially applied to benthic foraminifera is generally known to be more difficult and suggested to require more rigorous cleaning. One possible reason is that specimens from the Indian Ocean calibrations have had significantly lower contamination prior to the cleaning methodology. Another explanation is if the foraminifera tests from the Indian Ocean have been deposited in sediments of lower silicate.

Overall, the figures are well-made and easy to understand. I'm unconvinced this is a good idea, but if you (very lightly) shaded the "contaminated zone" above the peach-colored line e.g., in Figure 6, would it help drive home that almost everything's contaminated, or would it just add clutter?

Adding a feature in the graphs to help convey the extent of contamination is a good idea.

Line by line comments:

We thank the reviewer for the very thorough assessment of our manuscript. I relation to the minor revisions comments, we will address all purely editorial issues raised in the list below and correct the text accordingly. Additional comments are added below.

Lines 24-31: This paragraph seems out of place; reading it, I thought this paper was about to go in a very different direction. I think you could start from line 32 and be fine.

Lines 94-97: This whole paragraph or a statement of this kind belongs further towards the beginning – maybe at the end of Section 1?

Line 121 (?): What is the small inset panel on the right? I assume it's ship tracks but the information would be good to have in the caption.

Thanks for this comment. The inset panel displays the positioning of the T and S profiles within the GLODAPv2 database. We will clarify this in the next version of the manuscript.

Lines 153-216: Can you divide Experiments 1-3 into their own subheadings? I.e., "2.2.1. Preparation experiment 1: XYZ?" As it is now, it's a massive section that's difficult to follow. I could also see this being divided up where you explain the experiments simply and clearly in Section 2.2 and describe your findings in the Results.

We are grateful for your comment and include it in the revised version of the manuscript.

Line 219: There's inconsistency in foraminifera abbreviations: G. ruber vs. Cibicidoides wuellerstorfi, for example. I would go with the abbreviation, but be consistent either way.

Line 265: Where's panel D?

Thanks for spotting this error. We will correct it.

Line 325: Capitalize “c” in “c. wuellerstorfi.”

Line 369: “Figure ???11”

Section 4.2. header: Not sure why this is blue?

Lines 467-472: Is there any support for this in the literature or is it speculation? I hate to ask, but any possibility of elemental mapping to support?

Thanks for these comments. Elemental scanning is unfortunately beyond the scope of the project, but we will add references supporting the habit statements.

Lines 478-479: Missing parentheses.

Lines 484-493: I think Section 4.8 belongs in the introduction – it’s critical motivation for your cleaning tests but you don’t bring it up until the end.

Lines 495-521: This is too long of a block of text, it’s dense and hard to follow.

Lines 523-535: This is a great conclusion and summary! Bring some of this clarity to the introduction and method explanations.

Thanks for this comment. It is much appreciated.