Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-337-RC2, 2019

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Freshwater pearl mussels from northern Sweden serve as long-term, high-resolution stream water isotope recorders" *by* Bernd R. Schöne et al.

Anonymous Referee #2

Received and published: 17 September 2019

I have the same comments in a formatted PDF attached - the formatting gets lost in this online text input. General thoughts: Overall I think this manuscript has some very solid data, generated with methods that seem reasonable and appropriate. It is in general clear and well-written at the paragraph level (though see some suggestions below for improving clarity in a few places), though I did find myself circling back a few times to try to trace how certain calculations were made. Because of the complexity of their data sources (incomplete water δ18O values, incomplete air temperature records, etc.), the authors had to make some conversion/calculations, which I found a bit challenging to follow. More importantly (assuming I've understood what was done accurately) I have some broader concerns about some of the assumptions made

when transforming the data, and some concerns about the statistical approaches employed. I anticipate that addressing these concerns will not have major impacts on the conclusions reached, but I cannot be sure at this point. I am also not an expert on the North Atlantic, or Scandinavia, and I hope this manuscript receives review from someone with more familiarity on the climate systems discussed here. Overall I would consider these major revisions, as they do have the potential to change the conclusions. Broader Issues: âĂć Assumption sensitivity- The authors make several reasonable, but compounding and potentially important assumptions that should be further tested: o The authors provide a fairly straightforward equation (1) linking air temperature to water temperature, basically with a damping effect, based on the authors' previous research. This seems reasonable, but as this is an empirical relationship (that looks pretty good in original manuscript), it should really have some form of uncertainty estimates on the slope and intercept, and these uncertainties should be propagated to later equations. Otherwise it is difficult for the reader to assess whether the nature of this relationship is important later. o Another assumption the authors make (first stated in Line 69) is the use of annual increment width to determine weighted annual water temperature. On its face, this approach also seems reasonable, but again, the importance of the choices made here (with respect to weighting) is not evaluated, nor are uncertainties in SGI measurements. The equation in question (3) is also presented with no uncertainty estimates, as above – this could be important, especially given the fair amount of scatter in Figure 3b. There is clearly a meaningful relationship, but that does not imply that using SGI for predicting Tw* can be done without uncertainty, and there is certainly more scatter than the Ta-Tw relationship. o The use of annual weighting scheme (detailed in Table 2) is also reasonable, but is of course a necessary simplification that includes several other assumptions. Does the exact 143 day growing window matter? How much does the exact weighting of samples matter? Maybe a comparison to an unweighted calculation may be appropriate as a test for the importance of this step? o Basically – Tw is determined from Ta without uncertainty. δ18Os* is calculated using assumptions outlined in the point above. Then Tw is used along with δ18Os* to calculate δ18Owr* via an equation with no uncertainty (though none is presented in the original text to be fair). I'd like to see some effort to either propagate uncertainty, or to at least test the sensitivity of their final result to these assumptions. âĂć Statistics – I have some issues with the way statistics are handled in the manuscript. Like the assumptions above, it is quite possible that improving them will not change the ultimate conclusions of the manuscript, but there is certainly a possibility that some ideas will have to be revisited. Specifically: o One principle for using p-values is that it is not appropriate to run a large number of p-value tests and the use 0.05 as a threshold for significance. There a number of ways to account for this (like a Bonferroni correction) that provide a more appropriate significant threshold when running multiple tests. o That said, there are likely other multivariate statistical approaches that would be more appropriate to use for a variety of the questions the authors propose. They can probably get by without them here, but it might something to explore. o The comparison of δ18Ow and δ18Ow* (Fig. 6a) is determined to be "good" but by what standard. Can this comparison be improved/quantified statistically? Technical Comments: Broadly – if space isn't a concern, a table of all the various δasterisk/subscript variations (and potentially other abbreviations/nomenclature) and ideally even their method of calculation might be useful as a reference for the reader. I found myself constantly flipping back and forth to remind myself what each permutation meant. Abstract: I thought this was well written Line 35: "required" – could a quantitative measure of the required resolution be listed here? Also, this sentence would be better written a required resolution for . . .. what? added. Line 41: "short", "small" both could also have a quantitative measure added. Line 42: "signatures" at least according to Sharp (Principles of Stable Isotope Geochemistry) the word 'signatures' should be reserved for large reservoirs with consistent isotopic values (e.g. the ocean, the mantle). There are few other places below I will flag the isotopic terminology as a place to make slight (possibly pedantic) improvements. Line 88: I believe that typically for endangered species, if there was an approved permit for collection, that the permit should be referenced here and/or in the

acknowledgements, even if they were originally collected as part of another project. I am not sure if the journal has a policy on this issue though. Line 94: should this be "flow-through" lakes? Lines 95-105: I really appreciate the detail in here for sample preparation. Line 95: "dried from air" would be better written as "air-dried," or "dried in ambient air," or something similar; "from air" is awkward. Line 101-102: This mirrored idea is intriguing, but I am having trouble visualizing it. Given its potentially utility to others, a supplementary image of this set up would be nice. Lines 114-117: Given the importance of this method to this manuscript, it might be worth providing a few more details here, despite the previous publications. Lines 122-125: The text here seems to indicate they micromilled shell material by hand while maintaining consistent sample spacing. That's really quite challenging in my experience (though not impossible). It would be useful in Figure 2 to include an image of the actual drilled sample pathways. Was any attempt made to control (or later check) the depth of drilling? This would could be a significant source of variability. Lines 135+: I couldn't find access to a full table of data, but such a table must be provided with the final submission. That table (or tables) should include both the uncorrected δ18O value, but also the corrected δ18O for aragonite. Line 227: Why do we care that omitting GJ samples slightly increases R2? Line 228: I feel like a reference was omitted at the end of this line? Line 240: "Major common period" is not a calculation I am familiar with. Is this a formal calculation or a general observation by the authors? Line 281: "isotopes" should probably be "isotope values" or just "δ18O" I do not make more line-by-line comments to the text at this point, because I think the statistics and assumptions really need to be improved before the discussion is validated. However, here are some comments on the figures/tables: Figure 1: Potentially color/symbol differently for bivalve vs. water vs. temperature measuring sites? I'd propose removing the lines on 1b that presumably indicate districts within the province? They are not mentioned and visually complicate the figure. Moving Lat/Long markers outside 1b (especially the internal tickmarks) would also help simplify the figure. Potentially indicate the precipitation direction for NAO+/- on 1a if possible? Figure 2: Possibly include an image of the sample after drilling here as well? Figure 3: if more space is available, making 3a longer (extending x-axis) would be helpful in visually comparing the two records. Uncertainty on 3b best fit equation should be calculated and used (see comments above). Figure 4: Are these four just the best selections, or are they the most highly resolved? What was selection criteria? Like Figure 3a, extending these laterally would make viewing easier. Figure 5: A little small at present size overall as well. Could enhance clarity by labeling the rows and columns on the outside as well. Figure 6: Line 866- is there a statistical way to evaluate the "good agreement" observed here. Visually here, I do not find this as convincing as some of the other comparisons made. Figure 7: "δ18C" typo, also pretty hard to see at presented size. Table 1: Weird formatting issue (floating 800 in top right), line 799 – I believe the second "L" here should be an "i"? Table 3: see comments above about statistics and p-values. Also, probably don't need to report R and R2 – if R is reported just to indicate is relationship is positive, why not also instead provide the slope, which is more informative? I am a little torn here – I expect these relationships are pretty solid, but is running a regression on n=4, appropriate? Table 4: See comments above about p-values, and R and R2.

Please also note the supplement to this comment:
https://www.hydrol-earth-syst-sci-discuss.net/hess-2019-337/hess-2019-337-RC2-supplement.pdf

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-337, 2019.