Comment on cp-2021-142
Niels de Winter

In their manuscript, Johnson and colleagues present a nice dataset of seasonally resolved stable isotope transects through fossil bivalves to reconstruct seasonality during the mid-Piacenzian warm period. The study is timely, well thought-out and very relevant for the readers of Climate of the Past.

I enjoyed reading about the combination of this extensive, high-quality dataset and would like to compliment the authors on bringing the data together to give the reader an overview of the seasonality reconstructions from different individuals (Figure 8). The images of the *Aequipecten* shells (Figure 1) and the overview of the stratigraphic context of the shells (Fig. 3, section 3) are also a very useful addition to the field! On reading through the manuscript, I did encounter some aspects of the discussion which may require a bit more attention, or with which I did not fully agree, and I wanted to highlight these below so the authors could consider them in their revision. These comments are meant to improve the discussion of the nice dataset that is presented, which by itself is already a very valuable contribution to the field and certainly merits publication.

Preservation

The authors acknowledge that no preservation screening was done on the shell material (lines 389-391). In most deep time (pre-Quaternary) sclerochronological studies, I would consider such an investigation essential to demonstrate the reliability of isotope records. Just a few trace element analyses to test against incorporation of Mn or Fe during diagenesis (see Brand and Veizer, 1980), XRD profiles to test original aragonite preservation in the aragonitic species and/or SEM images to demonstrate original shell structure preservation would lend more confidence to the interpretations in the manuscript. That said, the authors do cite evidence of good preservation of specimens from the same or time-equivalent deposits and I know from personal experience that the preservation of these shells from the Lillo formation show excellent preservation, so I would not consider the lack of preservation screening in this study to be a big obstacle to interpretation of the results.
Transfer functions

In the manuscript, the authors nicely discuss the effect of applying several different transfer functions for the d18O-temperature relationship and a range of potential d18O values of the sea water on their d18O curves. Overall, I think this discussion is very honest and useful in showing the uncertainty on these d18O-based reconstructions, however I do not agree with the notion that the validity of transfer functions can be rejected or supported based on the data (e.g. lines 489-491; lines 679-684). In my opinion, the validity of proxy transfer functions like those for d18O can only be tested using modern carbonates precipitated at (approximately) known temperatures. Inferring the correctness of a transfer function based on the “fit” of fossil data with expected temperature outcomes runs the risk of circular reasoning. The discussion in lines 659-684, where outcomes of the d18O-temperature seasonality are compared with temperature reconstructions from ostracod and dinoflagellate assemblages is especially problematic, since the authors later (rightfully) argue that such assemblage-based reconstructions may be subject to bias (lines 926-929). My suggestion would be that the authors present the range of temperature seasonality outcomes they obtain from their fossil d18O data using various transfer functions and d18O values of the sea water as an uncertainty range. It is of course fine to discuss which outcomes fit better with previous reconstructions (which have their own uncertainty), but to conclude from these comparisons which transfer functions are best seems to push the interpretation a bit too far.

Statistics

In places where the uncertainty of the data is assessed (e.g. line 471-472) or comparisons between different records are made (e.g. line 570-574), the manuscript could benefit from more detailed statistical evaluation. For example, it would be more transparent if the measured values of the isotope standards are provided in a supplement and the actual mean value and standard deviation on these measurements is given in the text (line 471-472). In descriptions of the records, terms like “noise” (e.g. line 525) should be better defined and perhaps quantified. Statements like “substantially less variation” and “moderate positive covariation” (lines 570-574) should be backed up with statistical tests and quantification of uncertainty. Finally, I think the discussion would benefit from statistical evaluation of the seasonality outcomes and their uncertainty. The comparison between temperature reconstructions, on which much of the discussion is based, is heavily dependent on the way in which seasonality is calculated and the degree by which differences between reconstructions are statistically significant. The authors discuss how their method for extracting seasonality from the extreme values of d18O records influences the outcome (e.g. section 5.1), but the study design using a large number of specimens (data in Fig. 8 and Table 2 and 3) should make it possible to calculate ranges and uncertainties for summer and winter temperatures, which can be used to test statistically if some species or combinations of assumed d18O of seawater and transfer functions are in agreement with previous temperature estimates (see paragraph above).

Minor comments:

Line 128-129: In some species (e.g. Crassostrea gigas), shell sections in early ontogeny have been shown to by precipitated out of isotopic equilibrium (e.g. Huyghe et al., 2021), so this may not always be the best part of the shell to target for reconstructions.

Line 216-267: I really enjoyed reading this thorough review of the southern North Sea stratigraphy. I wonder if it would be beneficial to the reader to add rough paleo-depth
curves to the sections in Fig. 3 to make the evolution of the paleoenvironment in these different areas easier to follow.

Line 405-407: Does this penetration of the resin into shell affect the isotope analyses?

Line 411: Figure 3 does not show the drilling of *A. opercularis*, but instead shows stratigraphy of the mPWP sections. Perhaps this should refer to Fig. 1? (although this figure also does not show the drill holes)

Line 480-482: The authors should briefly explain here why the global d18Osw values are rejected here.

Line 489-491: See major comment above: The authors should explain why the Kim and O’Neil equation temperatures are “too low”. I would be careful with this type of reasoning about transfer functions based on the “expected” temperature value.

Line 546: Provide a number for “a great deal” to quantify the difference in growth rate.

Line 723: “overestimates” should be “overestimated”

Line 750-752: See also major comment about the transfer function discussion: I wonder if this reasoning about the height of the stratification factor based on the temperature outcome and its comparison with modern temperatures is not sensitive to circular reasoning issues.

Line 791-792: See comment above: I think one can almost never test the accuracy of proxy transfer functions (or the validity of d18Osw assumptions) based on their outcome on fossil data. This type of discussion requires independent evidence and/or modern calibration studies.

Line 926-929: If the assumption of ecological uniformitarianism does not always hold (with which I agree), the authors should be careful with their conclusions from comparison of temperature reconstructions with the outcome of ostracod and dinoflagellate assemblage studies elsewhere in the discussion.

References

Brand, U. and Veizer, J.: Chemical diagenesis of a multicomponent carbonate system–1: Trace elements, 50, 1219–1236, 1980.

Huyghe, D., Daëron, M., de Rafelis, M., Blamart, D., Sébilo, M., Paulet, Y.-M., and Lartaud, F.: Clumped isotopes in modern marine bivalves, Geochimica et Cosmochimica Acta, https://doi.org/10.1016/j.gca.2021.09.019, 2021.