Dear Geoscientific Model Development Editorial Board, Dear reviewers,

I would like to thank both reviewers for their comments and the editor dr. Olivier Marti for his moderation of the review process of my manuscript. I am delighted to read that both reviewers react positively to the revisions made to my manuscript and model code over the last review round, and that the manuscript is now ready for publication after minor revisions. Below, I offer a point-by-point reply to the questions and comments raised by the reviewers. In addition, I include a revised version of my manuscript in which I have tracked my changes. In addition to the suggested changes, I now refer to the Judd et al. (2018) model as the "GRATAISS" model throughout the manuscript text ("Growth Rate and Temporal Alignment of Isotopic Serial Samples"; following Ivany and Judd, 2022) in addition to other minor textual changes which I encountered on re-reading the manuscript.

Reviewer #1
Rereview of *ShellChron 0.4.0: A new tool for constructing chronologies in accretionary carbonate archives from stable oxygen isotope profiles* by de Winter

Overall, I am largely satisfied with the revisions made by de Winter to *ShellChron 0.4.0: A new tool for constructing chronologies in accretionary carbonate archives from stable oxygen isotope profiles*. Below I highlight four general comments, as well as some specific editorial comments. Pending minor- to moderate revisions, I believe this manuscript will be suitable for publication in *Geoscientific Model Development*.

I would like to thank Referee #1 for thoroughly and fairly re-reviewing my manuscript and model, and especially for their patience during the lengthy review process. As mentioned in reply to the previous round of comments, I am greatly indebted to the reviewer for their constructive criticisms on the first manuscript version which helped me improve both the manuscript and model to their current state. I am glad that my revision is positively received and grateful for the additional comments that helped me to finetune the manuscript for publication.

General comments

• **Fidelity of the examples**: My original main concern regarding the attenuation of the seasonal amplitude in the Case 1 and TEXEL examples seems to be resolved in this revision. However, the *Crassostrea gigas* example still shows some odd solutions (attenuation in Year 1 and multiple vertical lines in Year 2 in Figure 7B) not present in any of the other examples, that should be addressed.

Model accuracy is lower at the beginning and end of records, because the result in these places is based on a lower number of overlapping model windows. The accuracy also diminishes at the beginning and end of each modelling window (producing the "offshoots" in **Fig. 7B**), which is resolved by weighting model results at beginning and end of model less heavily. The lower sampling resolution of the oyster in first year and uncertainty and MC approach together make this start of the record more sensitive to random model uncertainty attenuating the record. I added some discussion of these results to section 4.3.

I'm also still slightly concerned by the systematically season residuals in Case 1 (Figure 6C; which also don't seem to fully match the distribution in Figure 6B). This seems to imply that something (e.g., bounds on the sine parameters, weighting of the windows) may be inducing a nonrandom bias. I encourage the author to explore this a bit more to determine where and why such biases occur, so that future users of the model have a better sense of whether such biases might affect their data.

Fig 6A-B shows that the bias occurs at peaks and troughs where small uncertainties in the position of samples drives larger uncertainties in age model. This effect is also visible in the $\delta^{18}O_c$ model fits (**Fig. 6A**). From **Fig. 6B** it shows that a few of these peak and valley datapoints drive the uncertainty in the monthly distribution of age inaccuracy, which is not normally distributed (see histogram in **Fig. 6B**). The Δt distribution in **Fig. 6B** matches the distribution in **Fig. 6C**, clearly showing the occurrence of small outliers on the positive Δt scale matching the minima in the $\delta^{18}O_c$ curve (and also the maxima in some places). Note that the offset is barely significant and the mean dt is not more than ±10 days. Nevertheless, these issues are now briefly discussed in section 4.2.

• **Test 1 Example:** I'm a bit confused by the example – if each window is supposed to span at least one year (L243-245) and the input growth rate of Test 1 was sinusoidal with no growth cessation (Figure 2A), why do the x-axes in Figure 4A,B only span ~250 days? This seems to imply that the model is imposing a growth cessation and condensing growth into 250 days. Is this the case? If so, can you achieve a better fit by changing some of the starting parameters of the model? Please address this is the text. It would also be helpful to also the "known" (i.e., input) temperature and growth rate curves on these panels for comparison.

The model does not impose growth cessation (see **Fig. 6B**), but coarse sampling means that the growth rate sinusoid is not accurately resolved everywhere within the window. Also, the difference between the first (point 1) and last datapoint (point 11) is not 365 because the next datapoint (point 12), which is precisely 365 days from the first point in the window, belongs to the new year and is hence not included in the window. Note that the modelled temperature curve closely approximates the known temperature. Longer modelling times should result in better approximation of the growth rate sinusoid (increasing maxn, see lines 301-303 and Fig. 5), however uncertainty in single windows is smoothed out by the moving window approach and weighting of solutions at the edge of windows, so the solution of one window not an fair measure of model performance. I added some discussion of this example clarifying these results at the beginning of section 4.2.

• **The impact of cumulative offsets**: Can you speculate on why the date offsets in the TEXEL example are not centered at zero? Is this phase shift an artifact of the windowed approach? e.g., If you were to omit the data from Years 1-3 and instead just model Years 4-9, would the results be different and, if so, by how much? What I'm getting at here, is that while I fully appreciate the advantages of the window approach and the formation of a single continuous time series, it ends up largely ignoring the only *a priori* age information we have about age (i.e., the placement of year markers or identifiable peaks and troughs in the data, which ultimately provide an annual chronometer). Many seasonal paleoclimate studies now stack temporally resolved data onto a single seasonal cycle, rather than analyzing them in time series, to get a climatological average (e.g., Tierney et al., 2020; Judd et al., 2019). In this sense, the *a priori* identified year markers help to minimize the propagation of uncertainty across the age model (i.e., one spurious year doesn't result in compounding error). A major advantage of the approach presented here is that it generates a continuous record, but I'm interested in understanding if and how this may propagate error across the record. I'm not sure that there's an easy solution here, but I think that the sensitively of the results to cumulative offsets and the range over which the data is modelled (e.g., full dataset vs. only a few select years) should be addressed in the discussion.

There is no cumulative offset in the model. The offset in the Texel example does not increase over record (accumulating uncertainty, see also histogram in **Fig. 6E**). Each window is fully independent from the previous window. Relative age results of windows are only later combined into continuous age-depth model by adding whole numbers of years, so there is no way the result of previous windows or years affects the seasonal phase of the next window.
The offset in the Texel example results from the way the real and modelled age-depth records are aligned at the same reference point in time. The small negative offset of the first datapoint from zero was thereby projected on the entire record (equally, without seasonal bias). I decided against scaling the entire modelled age downward to reduce this offset to prevent circular reasoning, but I will discuss this effect in the revised manuscript. Note that this choice of defining the "zero" of the age model result does not play a role in records with unknown ages, as the age result can be anchored to growth cessations or extreme values in $\delta^{18}O_c$. If all age windows were to be treated individually (as in the GRATAISS model), this "age alignment" decision will need to be made multiple times, potentially introducing uncertainty on year-by-year comparisons.
Finally, I agree that there are lots of uses for stacking data along the seasonal cycle (which does not require a continuous age record) as well as for continuous multi-year records. The advantage of ShellChron is that it allows data to be stacked relative to the seasonal cycle as well as providing a continuous record spanning multiple years. I now highlight this point in section 5.

Lastly, can you explain how the offset from actual age was computed for the Judd et al. model? Because each year is run individually using this model, the year markers define the starting and ending point of discrete years – so theoretically offsets should never be greater than >1 year (as suggested in Figure 8B), provided that year markers are correctly defined. In this sense and building on what is discussed above, it would perhaps be useful to see the year-to-year offsets rather than (or in addition to) the cumulative offset.

To obtain cumulative age estimates from the Judd et al. model, which indeed treats years individually, I exported the age data from the model and detected transitions from one year over to the next. Individual years were linked by having the results of the next year "window" in the sequence starting in the next year (e.g. 365 days or more after the first datapoint in the current year), thereby retaining the seasonal phase of the datapoints and preventing age inversions. This procedure results in gaps of time being introduced in the age model result due to the results of two consecutive years not matching up perfectly, which accumulate along the record to result in offsets larger than 1 year (see **Fig. 8B**).
I concur that these gaps are not really a result of the Judd et al. model but are inevitable when using it to create multi-year age models from $\delta^{18}O_c$ records. This issue is one of the main reasons why I set out to make ShellChron, as most $\delta^{18}O_c$ records record multiple years and applying the Judd et al. model often results in overlapping ages for consecutive $\delta^{18}O_c$ measurements (age inversions) or large time gaps between model years. Note that this issue may not be visible when stacking multiple years along a seasonal cycle, as is often done in paleoclimate studies, but this data treatment approach does not solve the issue: Age inversions, which result from inaccuracies in age modelling, may still be present at the beginning and end of the years, but a stacking approach will make it harder to detect them. The advantage of ShellChron is that the uncertainties at the edges of year windows are quantified and included in the model result, thereby making it possible to estimate the confidence of the age model and the likelihood that age inversions occur in the record from overlapping errors on results for consecutive datapoints.
This comment by the reviewer highlights an important difference between the two models which I agree needs some discussion in the manuscript. I add the year-by-year age uncertainty of the GRATAISS model to **Fig. 8B** and include the discussion above into section 4.4 and 5 of the manuscript.

• **Speleothem example**: In my opinion, the speleothem example is neither necessary nor constructive. As the model is currently written, it is not designed for speleothems (L167-170) and the speleothem example only serves to highlight this. In most cases, speleothems violate the assumption of sinusoidal oscillation in δ18O and thus would likely benefit from a different age modelling approach. The manuscript is already quite long with several examples, and there's no need to add length by explaining why such systems are difficult to model (e.g., L477-496; L588-605).

I disagree with the statement that the speleothem example is not constructive in the manuscript. The example serves to highlight the limitations of the ShellChron model while at the same time showing how the model performs when pushed towards these limitations. The speleothem example thereby serves as a benchmark for a worst-case scenario when applying ShellChron on troublesome or poorly constrained $\delta^{18}O_c$ records. I do agree that most speleothems violate the assumption of sinusoidal temperature or $\delta^{18}O_w$ oscillations (sinusoidal $\delta^{18}O_c$ oscillations are not assumed in the model, as the modelled $\delta^{18}O_c$ is built up of sinusoidal temperature and skewed sinusoidal growth rate records). However, sinusoidal oscillations in $\delta^{18}O_c$ of the drip water feeding speleothems have been demonstrated in cave monitoring studies (e.g. Baldini et al., 2008; Van Rampelbergh et al., 2014; Vansteenberge et al., 2019; see discussion in section 4.3 and 5). This caveat allows the assumptions of ShellChron to be discussed (as pointed out by the referee) and the modular character of ShellChron to be showcased with an outlook towards future improvements. In summary, I prefer to leave the speleothem example in the manuscript, unless the editor has a strong objection to this.

Specific comments
L217: add comma between "included" and "which"
Done
L358: change "all." to "al."
Done
L444-446: awkward phrasing; consider revising to "The lower sampling resolution later in the record mutes this variability and further illustrates that..."
This has been rephrased accordingly
L468: "real" feels colloquial and nonscientific; I'd recommend changing "real" to "known", both in the text and the figures (e.g., "the "known" age of the samples in these natural carbonates is not truly known.")
Good comment, I agree and rephrased to "known" throughout the manuscript and the figures

L476: remove “very” (such descriptors are unquantifiable)
Agreed, “very” was removed.
L671: change “eliminate” to “minimize”
Agreed, this has been rephrased
Figure 2: change the blue bar to reflect the window used in the Figure 4 example.
The blue bar has been moved to the first growth year in the revised version, in accordance with the window used in Fig. 4.
Figure 7: for consistency with Figure 6, I suggest keeping the isotope residuals histogram in the upper left corner and the date residuals histogram in the lower right corner
A good suggestion, I replaced the histograms accordingly in the revised Fig. 7

New references cited

Judd, Emily J., Linda C. Ivany, Robert M. DeConto, Anna Ruth W. Halberstadt, Nicole M. Miklus, Christopher K. Junium, and Benjamin T. Uveges (2019) "Seasonally resolved proxy data from the Antarctic Peninsula support a heterogeneous middle Eocene Southern Ocean." *Paleoceanography and Paleoclimatology* 34(5).
Tierney, Jessica E., Christopher J. Poulsen, Isabel P. Montanez, Tripti Bhattacharya, Ran Feng, Heather L. Ford, Barbel Honisch et al. (2020). "Past climates inform our future." *Science* 370(6517).

Reviewer #3

My comments here are informed by the original ShellChron manuscript, two sets of reviews, the revised manuscript, and detailed author responses to the original reviews, all of which I received in order to provide this feedback.

ShellChron is introduced as an open source, R-based package with which to interrogate subannually resolved proxy data from records with nominally sinusoidal input variables. The package builds upon a series of earlier publications by other authors over two+ decades detailing computational approaches with which to recover seasonally resolved environmental parameters from proxy data archived in accretionary, mineralized, largely biogenic records. A series of model and real test cases are explored so as to evaluate the performance of this model with data sets of varying internal complexity. Advantages and limitations are discussed in some detail, along with the range of possible applications.

I found the first round of reviews of this manuscript to be thorough and detailed. They identified several specific shortcomings of the initial iteration of the model that led to spurious or systematically biased output, enabling the author to refine and improve the code. De Winter’s responses to reviewer feedback, as evidenced both in his written text and in the new rounds of model performance, are substantive, thoughtful, and thorough. Most important was the tendency of the model to underestimate observed seasonal range in the carbonate d18O values of test cases, a function in part of too-broad constraints on how much growth varies over a year (a problem identified also by Judd et al) and in how time was equated to distance along the accretionary axis. Careful examination of and adjustments to the code identified the sources of this problem and new simulations demonstrate that the issue appears to be well resolved. Test cases illustrate very good agreement between data and model. This is a nice package!

The major new advantage of ShellChron is that it combines the continuous moving window approach of earlier models with the intra-annually varying growth rate parameterized by the recent Judd et al. 2018 model for individual years of data and adds an approach for propagating error in both the X and Y axes through to the model output to generate a confidence envelope. As well, ShellChron expands availability of the Judd et al. approach from the original MATLAB to the R language, and its modular design allows for the user to parameterize a function describing variable d18O water and/or to use a seasonal growth model other than a skewed sinusoid. This, combined with the detailed and clear manuscript, all makes the model quite accessible to a wide audience and applicable to a range of situations. Since it incorporates and builds from previous work though, I would urge him to explicitly recommend citing both this paper/program and the Judd paper together. That way both the R package and the work that inspired it can receive adequate credit. Judd is an early career researcher, and we know how important citation rates are at that crucial career stage.

I would like to thank Referee #3 for his kind review of my revised manuscript and model, and for the suggestion. I fully agree that, while ShellChron is in many ways different from the GRATAISS model, the principles underlying it are very much derived from the model by Judd et al. I will therefore include the suggestion to cite this important work alongside this manuscript whenever ShellChron is used by the community. I added this suggestion at the end of the abstract where it will be most visible, but if the editor believes there is a more appropriate place for this statement, I would be happy to move it elsewhere.

Minor questions/edits:
Lines 258-260 (and elsewhere) – should this be 'differential weighting', rather than 'weighing' in this discussion?
Agreed, I rephrased "weighing" to "weighting" throughout the manuscript
Why in Figure 4A and 4B are full sinusoids not depicted?
The full model result (i.e. the age estimate for each of the datapoints in the example window) is depicted in these figures. See my reply to a comment by Referee #1 for an explanation of why the difference between the age of the first and last datapoint in the window is less than 365 years.
Line 339 – some text missing?
No text is missing here. Instead, the sentence starting with "in which…" explains the meaning of the symbols in the above formula 10.
Line 358 – should be 'et al.'
This is rephrased accordingly
Line 418 – 'with in'
Line 418 already contains the phrasing "with in". It is not clear to me if the reviewer suggests any rephrasing is needed here.
Figure 6C and F need a bit more description and qualification to make them clear. As well, variation in 6C looks to be (nearly) significant over the year, but this is minimized in the text. Worth more discussion?
I added more detailed description of Fig. 6C and F in the caption. The age offsets in these figures are indeed significant in some monthly time bins. These instances are now discussed in the manuscript text. See also my reply to comments by Referee #1 on these figures.
Line 566 – 'within' should be 'with'
Agreed, this is rephrased accordingly

References

Baldini, J. U. L., McDermott, F., Hoffmann, D. L., Richards, D. A., and Clipson, N.: Very high-frequency and seasonal cave atmosphere PCO2 variability: Implications for stalagmite growth and oxygen isotope-based paleoclimate records, Earth and Planetary Science Letters, 272, 118–129, https://doi.org/10.1016/j.epsl.2008.04.031, 2008.

Van Rampelbergh, M., Verheyden, S., Allan, M., Quinif, Y., Keppens, E., and Claeys, P.: Seasonal variations recorded in cave monitoring results and a 10 year monthly resolved speleothem δ18O and δ13C record from the Han-sur-Lesse cave, Belgium, 10, 1821–1856, 2014.

Vansteenberge, S., Winter, N. de, Sinnesael, M., Verheyden, S., Goderis, S., Malderen, S. J. M. V., Vanhaecke, F., and Claeys, P.: Reconstructing seasonality through stable isotope and trace element analysis of the Proserpine stalagmite, Han-sur-Lesse Cave, Belgium: indications for climate-driven changes during the last 400 years, 1–32, https://doi.org/10.5194/cp-2019-78, 2019.

Ivany, L. C. and Judd, E. J.: Deciphering Temperature Seasonality in Earth's Ancient Oceans, 50, 123-152, https://doi.org/10.1146/annurev-earth-032320-095156, 2022.