Response to RC1: 'Comment on gchron-2023-22', Maarten Blaauw, 27-Sep-2023

[1,2,*]Robin B. Trayler [3]Stephen R. Meyers [4]Bradley B. Sageman

[2]Mark D. Schmitz

[1]Department of Life and Environmental Sciences, University of California, Merced, CA

[2]Department of Geosciences, Boise State University, Boise ID

[3]Department of Geosciences, University of Wisconsin, Madison, WI

[4]Department of Earth and Planetary Sciences, Northwestern University, Evanston, IL

[*]Corresponding author: rtrayler@ucmerced.edu

- This manuscript proposes to combine proxy data of orbital forcing with radiometric dates in order to produce integrated age-depth models from cores. It builds a piece-wise linear model with constrained accumulation rates and the possibility of hiatuses, and treats the response to the orbital forcing as a constant offset 'a' in years for each section. The method is tested using synthetic and real-world data, and shows huge enhancements in precision compared to dates-only age-models (but see below).

- Generally the method is described well and placed in the wider context through an interesting review of existing methods. However, I would like to see some more detail on how the parameters are estimated, how the priors are set and how these settings affect the age-depth models (robustness analysis).

We thank Dr. Blaauw for their complement about the clarity of our study.

- Section 3.2.3, what limits are put on the accumulation rate, and why are you using a uniform distribution? Why not use a more informative prior on accumulation rate, such as a gamma with a specified mean and shape (the latter can be put at very permissive values, e.g., 1.1)?

We choose to use a uniform distribution as a vague prior distribution so that the cyclostratigraphic likelihoods are the primary control on sedimentation rate. The joint inversion with the radioisotopic dates further limits the possible accumulation rate, and the dates themselves can be used to estimate appropriate bounds for the uniform prior distribution. For example, calculating the slope between each date-pair in sequence gives an average sedimentation rate. Making the same calculation at say ±5σ from the young-tail to the older-tail of a date-pair gives "worst case scenario" sedimentation rates that can be used as bounds for the uniform prior.

- This would then also diminish the likelihood of cycles that require extreme accumulation rates (i.e., the harmonic analysis of Fig. 2c/f would show darker colours for less realistic accumulation rates).

  Just to clarify, Panels 2C and 2F are not showing the probability of accumulation rates. Instead they are showing spectral amplitude, calculated over a moving window. These plots are used to interpret stratigraphic "layers" were sedimentation rate is stable but do not inform the absolute value of sedimentation rate necessarily.

- Are there limits on the hiatus size as well, and is the prior also uniform or rather gamma as suggested by Fig. 7?

  We did not define a prior distribution for hiatus duration, except for the limitation that hiatus durations be positive values, such that they cannot violate superposition. The probability of a hiatus duration is estimated in the same manner as the other model parameters (anchor age, sedimentation rate(s)).

- The difference in modeled precision between `BChron` and astroBayes is huge, especially in the case where a core has only few radiometric dates. That said, how robust is the assumption of linear accumulation over long time-scales (e.g., the long section of Fig. 3b)? Although this is discussed, I still find it hard to believe that a geological sedimentation process really was exactly linear over large amounts of time - if this assumption is not met, then the reconstructed precision will be illusionary high.

Dr. Blaauw is correct that the increase in precision of `astroBayes` compared to `BChron` results from our choices of a much simpler sedimentation model. We do feel that this choice is justified however, since the stratigraphic information about sedimentation rate that the astrochronology provides, is not trivial and shows that sediment accumulation really can be near-linear for very long periods of time. For example in Figure 5B, it's quite possible to draw a vertical, straight line through the highest amplitude frequency track (~1 cycle/m) within each of the layers we have defined. So while at the very fine scale sedimentation is absolutely a variable process, the astrochronology does show that it can be approximated by a series of linear segments. Clearly this will not always be the case and the suitability of using `astroBayes` to model different datasets will need to be assessed on a case by case basis.

Ultimately our goal is to capture the "true" age model within the `astroBayes` posterior even if we are somewhat simplifying the problem. For example, in figure 2C, the second layer from the base of the section has a varying sedimentation rate that is only partially approximated by our choice of treating it as a single layer. Nevertheless, inspecting the age-depth models in figure 3A-D shows that even when our assumptions of more-or-less constant accumulation are violated the true age-depth model still falls within the 95% credible interval of the posterior, which is reproduced in nearly all cases (see line 300).

- The proposed method uses a limited number of sections of linear accumulation, e.g. <10. Does increasing this to say 50 or 100 shorter linear sections affect the age-depth models by much (e.g., in terms of smoothness and reconstructed uncertainties)? Could you explain a bit more how the 'elbows' (z) between the sections are chosen and how they can be made to vary over depth? This model reminds me of Bpeat (Blaauw and Christen 2005 Applied Statistics 54, 805-816), which modeled accumulation using a handful of linear sections and where the depths of the 'elbows' were part of the parameters to be estimated (this also included hiatuses and outliers).

Currently layer thickness is somewhat constrained by how the probability calculations are made (See line 320 in submitted manuscript).

> "… a limitation of our model is that each layer must contain enough time and astrochronologic data to resolve the astronomical frequencies (f) of interest."

Since the layers must contain a a sufficient amount amount of cyclostratigraphic data, increasing the number of model layers degrades the likelihood calculated using equation 2. The cyclostratigraphic sampling rate also plays a role here as the Nyquist frequency within each layer must be high enough to capture each of the target frequencies. Relaxing the first constraint is possible but would require an entirely new modeling framework. Practically this would require a re-write of the core `astroBayes` function and would also greatly increase the computational time required.

It was only briefly mentioned (line 185) in the manuscript, but the implementation in the `astroBayes` package does allow the user to assign uniform uncertainties to the layer boundary positions. The model will randomly adjust the boundary position within the uncertainty for each iteration so that the initial choice of layer position is somewhat less important. We also note that this is the same basic approach that Malinverno et al. (2010) took to deal with constructing Bayesian floating age models from cyclostratigraphic data.

We have expanded part of the *Model Construction* section (lines 176 - 186) so it now reads:

> "The selection of layer boundary-positions is an important user defined step, that is informed by detailed investigation of the cyclostratigraphic data. Evolutive harmonic analysis (EHA) is a time-frequency method that can identify changes in accumulation rate by tracking the apparent spatial drift of astronomical frequencies. Expressed as cycles/depth, high amplitude cycles may"drift" towards higher or lower spatial frequencies throughout the stratigraphic record. Assuming these spatial frequencies reflect relatively stable astronomical periodicities, the most likely explanation of those spatial shifts is therefore stratigraphic changes in sedimentation rate (Meyers et al., 2001). That is, stability in spatial frequencies reflects stability in sedimentation rate,

and show that in these cases sedimentation can be approximated by a small number of piecewise linear segments.

We visually inspected EHA plots to develop simple sedimentation models (e.g., 2B) for our testing data sets. We choose layer boundary-positions (z_1 – z_i) by identifying regions with visually stable spatial frequencies. For example, in 2C, there is a continuous high-amplitude frequency-track between 2-4 cycles/m. Based on visual shifts in this frequency, we choose three layer boundaries, such that this frequency track can be approximated by a vertical line within each layer. In the computation implementation, we also allow the layer layer boundary-positions to vary randomly (within a user specified stratigraphic range) to account for stratigraphic uncertainties in boundary-positions that arise from the fidelity and our inspection of the of the data, similar to the Bayesian cyclostratigraphic approach of Malinverno et al. (2010)."

- What about a potential alternative model, also with set boundaries between the different known sections of nearly but not entirely linear accumulation (e.g., a bit like Bacon but with a very high and strong prior memory on accumulation rate, so a very low variability of accumulation rate over a section, but still some possibility of deviation from an entirely straight line), and with very permissive/wide prior accumulation rates for each of these sections so rates can jump from one z to the next?

This is a **great idea**, that we feel is outside the scope of the current study. Currently, implementing this modeling framework would require rewriting a large chunk of the core `astroBayes` code/functions. However, the model that Dr. Blaauw describes sounds a lot like a hybrid of `bacon` and `astroBayes`, and perhaps future development of the `astroBayes` package could implement an "astroBacon" modeling framework.

- Some more information on the MCMC settings and decisions could be helpful, e.g. in Supplementary Information. Perhaps also provide a short tutorial, much like on the helpful GitHub pages but using the examples of the manuscript and with more information as to what steps are taken and why.

This is a helpful comment, especially from a user-friendliness perspective. We will include the Github tutorial as a vignette in the `astroBayes` R package to make it more discoverable.

- Table 2: The estimates are given as 7 digits, which implies that the length of each of the periods is known to the month (!). Should the values not be rounded to a more realistic precision, and if so, what would that precision be (millennial I'd say)? Are there any estimates of the size and shape of the uncertainties related to the period/frequency estimates of the different orbital cycles?

  Dr. Blaauw is correct that the reported number of decimal places here do not reflect the true precision. There are uncertainties associated with each of the Milankovich frequencies, however, the periods have changed over very long-term time scales (e.g., tens of millions of years), so appropriate periods for say Eocene vs. Pleistocene records will be different. There are various tools available to calculate astronomical periods in deep time (see here) and also (Laskar et al., 2011).

- Does it matter for the harmonic analysis where in time each of the cycles of Table 2 starts?

  No, it does not matter. Since the cyclostratigraphic data is transformed from the stratigraphic position domain into the frequency domain, our method explicitly accounts for the phases of each cycle.

- Section 3.2.2, was no outlier analysis done?

  We thank the reviewer for this suggestion. We did not initially include an outlier analysis, but we have done so now. The sections below have been added to the manuscript in the *Testing and Validation* and *Results* sections respectively. The corresponding R scripts to reproduce the results have likewise been added to the robintrayler/astroBayes_manuscript Github repository.

Sensitivity To Outlier Ages

We also tested the sensitivity of `astroBayes` to the inclusion of outlier ages. We repeated the tests from section 3.3.2, with one additional step. After the generation of stratigraphically-randomly

distributed dates, we used Monte Carlo methods to select one date. This date was then randomly adjusted by ±1σ to ±4σ. This creates a date that is either broadly comparable with the underlying true age model (e.g., ±1σ to ±2σ), or outlier ages that may introduce stratigraphic miss-matches (e.g., ±3σ to ±4σ). We choose to introduce these more subtle outliers, since we feel more extreme outlier ages can often be identified and excluded *a priori* based on inspection of the radioisotopic data (Michel et al., 2016). We repeated this procedure 1,000 times using either 2, 4, 6, or 8 dates for each data set (as in the section above), so that 1/2, 1/4, 1/6, and 1/8 dates would be considered an outlier. Each simulation ran for 10,000 MCMC iterations with a 1,000 iteration "burn-in".

Model Validation

… `astroBayes` is somewhat sensitive to the inclusion of subtle outlier radioisotopic dates. The inclusion of outlier ages lowered the proportion of the true age-depth model that fell within the 95% credible interval of the `astroBayes` to 89% for TD1, and 88% for CIP2. The relative percentage of outlier ages also does not appear to have a strong influence. …

Details

- Line 13, 49, spatial-temporal? How does is the spatial component involved? Do you rather mean vertical, depth-scale resolution? Spatial could be interpreted as 'horizontal' resolution, as in, how representative is a core of wider spatial events. How would one define high temporal resolution? Rather, mention that this is at, e.g., 10^5-6 yr resolution.

  We have replaced the references to "spatial" with "stratigraphic" which is what we meant.

- 15, high-precision, quantify

  We added "(<±1%)", also see the comment below on the second use of this term on line 94.

- 94, again high-precision - this seems an unnecessary qualifier here as no-one would aim for low precision.

7

The term "high precision" is often used in deep-time geochronology to distinguish in-situ methods (LA-ICPMS, SIMS) from whole crystal methods. Individual spot analyses from in-situ methods commonly have a precision of ±3-5% and ~±1% precision for weighted means. In contrast, Modern CA-ID-TIMS (U-Pb) and multi-collector mass spectrometers ($^{40}Ar/^{39}Ar$) have a precision of <±1% for single crystal analyses and approach <±0.1% for weighted mean ages. Also see box 1 in (Schmitz and Kuiper, 2013).

- 270, what MCMC thinning was used?

We did not thin our Markov chains. Our understanding is that there continues to be a debate about whether thinning Markov chains is strictly statistically necessary or if mostly used to address computational / computer storage constraints when not thinning would generate unmanageably large vectors or matrices of data, with different studies supporting both conclusions (Link and Eaton, 2012; Owen, 2017). That said, MCMC thinning is a straightforward improvement that can be added to the model code, either simultaneous with model iterations (as in `Bchron::Bchronology()`) or applied post-hoc (as in Dr. Blaauw's own `rbacon::thinner()`), and this can be added to the development version (and integrated into a future version) of the `astroBayes` package.

- 299, The test was done using simulated sections of constant accumulation, so that the model closely follows the simulated truth is no surprise.

We agree that this is not a surprising result but would like to point out that neither of the testing data sets have "constant" accumulation. See the example of the ~10-15 meter layer in panel 5C where the evolutive harmonic analysis shows that sedimentation rate is gradually changing. Furthermore, the sentence this comment refers to is discussing how model results are very consistent, *even* when the number and stratigraphic position of dates is variable which we feel is an important result that would not necessarily be possible with other age-depth modeling frameworks. This is especially apparent in Figure 3 where the sedimentation rate is correctly estimated even in model layers that do not contain any dates, without the inflation in credible intervals seen with `Bchron`.

8

- 430, to evaluate > Fixed

References

Laskar, J., Fienga, A., Gastineau, M., and Manche, H.: La2010: A new orbital solution for the long-term motion of the Earth, Astronomy & Astrophysics, 532, A89, 2011.

Link, W. A. and Eaton, M. J.: On thinning of chains in MCMC, Methods in ecology and evolution, 3, 112–115, 2012.

Malinverno, A., Erba, E., and Herbert, T.: Orbital tuning as an inverse problem: Chronology of the early Aptian oceanic anoxic event 1a (Selli Level) in the Cismon APTICORE, Paleoceanography, 25, 2010.

Meyers, S. R., Sageman, B. B., and Hinnov, L. A.: Integrated Quantitative Stratigraphy of the Cenomanian-Turonian Bridge Creek Limestone Member Using Evolutive Harmonic Analysis and Stratigraphic Modeling, Journal of Sedimentary Research, 71, 628–644, 2001.

Michel, L. A., Schmitz, M. D., Tabor, N. J., Moontañez, I. P., and Davydov, V. I.: Reply to the comment on "Chronostratigraphy and paleoclimatology of the Lodève Basin, France: Evidence for a pan-tropical aridification event across the Carboniferous–Permian boundary" by Michel et al., (2015). Palaeogeography, Palaeoclimatology, Palaeoecology 430, 118–131, Palaeogeography Palaeoclimatology Palaeoecology, 441, 1000–1004, https://doi.org/doi: 10.1016/j.palaeo.2015.10.023, 2016.

Owen, A. B.: Statistically efficient thinning of a Markov chain sampler, Journal of Computational and Graphical Statistics, 26, 738–744, 2017.

Schmitz, M. D. and Kuiper, K. F.: High-Precision Geochronology, Elements, 9, 25–30, 2013.