Authors response to reviewer comments on the manuscript “Nonparametric estimation of age-depth models from sedimentological and stratigraphic information” (EGUSPHERE-2024-2857)

Line numbers correspond to the line numbers in the pdf with tracked changes.

We address comments from each reviewer individually. For this, we list the publicly visible author responses (https://egusphere.copernicus.org/preprints/2024/egusphere-2024-2857/#discussion) and discuss what changes were made in the revised manuscript. Reviewer comments are listed in red.

Comments by Jarno Huygh

Comment: ” The manuscript aims at providing stratigraphers with two non-parametric methods for constructing age-depth models, taking into account uncertainty and integration of a wide(r) range of information available to the researcher into the age-depth model. Both the manuscript and the supplementary material (R-code etc.) are very detailed and well organized. The quality of the included figures is good. This manuscript is targeted at stratigraphers and the following comments are from the perspective of a (cyclo)stratigrapher. No comments will be given on the selection of the case studies and the discussion thereof.

Even though the manuscript is aimed at stratigraphers, it is very mathematics-oriented, very abstract, and, requires a considerable statistical background and large time-investment to fully grasp. Figuring out how to (correctly) apply the developed methods (using the admtools package) would equally demand a very large time-investment. These factors create a very high barrier and may likely result in few researchers using the methods.

In addition, the manuscript fails to clearly highlight the added value of using these methods and thus to convince researchers to invest their time in learning how to apply them. Especially when working on Paleozoic cyclostratigraphic case studies (that inherently have very large time-domain uncertainties), the added value of the proposed methods ‘appears to be’ negligible. Lastly, age-depth models are generally **not** just crude simplifications (as is mentioned in the abstract) – in fact, a vast array of information, data and insight is combined by the researcher to construct them.

I would strongly recommend to drastically reduce the size of subchapter 2 (lines 90 to 341) and to remove all formulas to the supplementary material. In addition, I would recommend emphasizing why stratigraphers need these methods, when they are most useful and, what pitfalls to watch out for when usign them (use of plain language may be useful here). Lastly, this manuscript would, after publication, greatly benefit from an online course on how to correctly apply given methods. Clearly, great effort and talent was put into this work and it would be great to see it applied in research.”
Response: “Dear Jarno

Thank you for taking the time to read the manuscript and for your comments.

Thank you for your praise of the structure and detail of the manuscript and supplementary material such as the R code. We value code documentation and clarity to make methods and results as well documented as possible to facilitate methodological clarity.

One of the main points you raise is that the manuscript is very mathematics oriented ("Even though the manuscript is aimed at stratigraphers, it is very mathematics-oriented, very abstract[…]") and suggest moving most of the methods description into the supplementary materials ("I would strongly recommend to drastically reduce the size of subchapter 2 (lines 90 to 341) and to remove all formulas to the supplementary material.")

We believe that scientific methods should be documented in detail, so that the tool is not a "black box" to the users. In doing this, we follow the example of previous articles in this field, which served as guidance for us, including Blaauw & Christen (2011) or Haslett & Parnell (2008). It would be patronizing towards the readers to assume they cannot follow the derivation of the method. We support the expectation that one should aim at understanding the method one uses, and the premise of what we propose is in fact delegating more agency to the users in choosing the parameters and assumptions. Removing the equations would go against this intent. The brief summary of the underlying logic at the beginning of each section (lines 120 to 125 and 165 to 180) gives readers the opportunity to skip the technical parts of the manuscript. We leave the choice whether to move the equations to the supplementary materials or to keep them in the main text body to the editor.

You raise the point that the methods are hard to apply and will take time-investment to learn ("Figuring out how to (correctly) apply the developed methods (using the admtools package) would equally demand a very large time-investment"), and suggest that we should organize workshops on the methods ("this manuscript would, after publication, greatly benefit from an online course on how to correctly apply given methods.").

We intend to develop such a workshop, and are committed to improve the documentation and vignettes associated with the R package. We welcome all specific comments and suggestions for these components, and invite the scientific community to improve them or suggest improvements to them via GitHub according to the package contribution guidelines.

You state that the added value of the presented methods is not clear, "*especially when working on Paleozoic cyclostratigraphic case studies (that inherently have very large time-domain uncertainties)"*. We are sorry if this is not clear from the current version of the manuscript, and are happy to clarify this when revising it. However, considering that the lack of systematic reporting of uncertainty by cyclostratigraphic studies was one of the challenges listed in the seminal reproducibility review study by Sinnesael et al. (2019), we are surprised to hear that the added value of incorporating the uncertainty of sedimentation rate estimate is not clear. We believe that the presented methods provide the opportunity to include previously unused data (e.g., semi-quantitative estimates of sedimentation rates, knowledge on tracer fluxes) into the construction of age-depth models. And we stand by our claim that admtools allows incorporating uncertainty from evolutive cyclostratigraphic methods, which is otherwise nowhere exploited in current literature."

Changes made:

- Phrasing of the abstract that age-depth models are simplifications was removed (lines 3 & 4), language throughout the manuscript was adjusted.
- A workshop on the developed methods will be given at the CycloNet meeting 2025 in Utrecht, and workshop materials will be made publicly available (https://www.uu.nl/en/research/department-of-earth-sciences/cyclonet-meeting-utrecht)
- We have left the decision to move the mathematical part to the supplementary material to the editor, who did not make this request. We have decided to keep all mathematical derivations in the main text body, as we maintain that methodology should be thoroughly documented and is a key part of the publication.
- Added value of the methods and potential sources of information were added to the introduction (lines 63 to 69, 94 to 98, 655 to 676)

Comments by Richard Zeebe

Comment: "The manuscript discusses age-depth models in general, as well as age models for specific events, including the Paleocene-Eocene Thermal maximum (PETM). Unfortunately, the omission of references to immediately relevant work on the subject and the lack of discussion of the authors' results in relation to previous work is highly problematic.

The authors' PETM age model uses 3He data from Farley and Eltgroth (2003). Omission of the immediately related study by Murphy et al. (2010) using 3He for the PETM is hence rather incomprehensible:

Murphy, B. H., K. A. Farley, and J. C. Zachos. An extraterrestrial 3He-based timescale for the Paleocene-Eocene Thermal Maximum (PETM) from Walvis Ridge, IODP Site 1266. Geochimica et Cosmochimica Acta 74, no. 17 (2010): 5098-5108.

Murphy et al. (2010) designated the end of the PETM CIE recovery to occur at 217 (+44/-31) kyr, a much longer time interval than suggested in the present manuscript (Fig. 6).

Zeebe and Lourens (2019) derived a PETM duration of 170 (+-30) kyr from onset to recovery inflection, again a much longer time interval - and at odds with the present manuscript.

Zeebe, R.E. and Lourens, L.J., 2019. Solar System chaos and the Paleocene-Eocene boundary age constrained by geology and astronomy. Science, 365(6456), pp.926-929.

Regrettably, none of the studies above is mentioned or referenced in the ms (there are more). Neither are the large discrepancies discussed relative to previous work. Perplexingly, there is also no reference to any of the PETM landmark papers led by Zachos in the reference list.

Given the above, the manuscript should not be considered for publication without an in-depth analysis and discussion of relevant previous work on the PETM and PETM duration, most notably those studies at odds with the present manuscript, including the references mentioned above.

Richard E. Zeebe"

Response:" Dear Richard

Thank you for your comment and feedback on this manuscript.

The aim of the PETM example was to examine how deviations from the assumptions of constant 3He flux change the inferred age estimates. The assumption of constant flux is an obvious weakness of previous age-depth models as acknowledged in the respective publications, which is why we chose this example to demonstrate the flexibility of FAM. We provide all code and documentation for this example, ensuring it is transparent and reproducible, and show that short-term variability has only a minor effect on the estimated ages.

We apologize if this example was perceived as problematic – as stated in the manuscript, our intention was not to determine new durations for the PETM, but rather explore the robustness of the model assumptions used in previous age-depth models (lines 450f). We are happy to adjust the phrasing of the results and problem in the abstract and the example, and include more references for context on the discussion on the PETM duration. As preliminary analyses, we applied FAM to the 3He record from ODP site 1266 presented by Murphy et al (2010) (https://doi.org/10.1016/j.gca.2010.03.039) and arrive at very similar durations as they do (e.g., duration clay layer: Murphy et al. (2010): 167 kyr, our results: 170 kyr). If both you and the editor agree that the example from site 1266 provides a more suitable example to demonstrate the methodology and adds to the manuscripts clarity and reception, we would be happy to substituting site 690 with site 1266."

Changes made: We have changed the example on the PETM, and now estimate the age-depth model using the data from Murphy et al. (2010). The updated analysis can be found on lines 498 to 565 (section 3.2), and results on the PETM duration align with those of the mentioned manuscripts (Zeebe and Lourens 2019) and other studies (Sluijs et al. 2007). In the example, we provided more context on the effect of varying Helium fluxes on age estimates and the PETM in general.
The phrasing in the abstract (line 21) and the example was clarified to emphasize the aim of the example – examining the effect of short-term fluctuations of Helium fluxes on age estimates (lines 510 to 511).

Comments by Claire O. Harrigan

Comment: "One thing I appreciated about this work is the effort that the authors put into writing clear and direct sentences. They generally did a good job explaining why they did a certain step in a process. I think they are capable of bringing that level of clarity and intentionality to all parts of the manuscript, and I would like to see some additional work around stating and justifying some of their assumptions. It is clear that stating assumptions is a key motivation for developing the methods they describe, so it was slightly jarring when coming across an assumption they made (or a statement about assumptions) that did not feel fully explained. For example, I had some unease around the suggestion that these methods produce age-depth models that are "data-driven rather than assumption-driven" (lines 69-70). These methods clearly incorporate assumptions (e.g., lines 91, 392, 459-460) and data, as all models do. It seemed like the implication was that these methods were less assumption-based than other models, and I am not sure that is entirely true. In particular, I found the assumption of positive sedimentation, and by implication, sedimentation with no hiatuses or erosion (e.g., lines 91, 286-288), a potential problem with this work that could lead to FAM and ICON producing results that are not a strictly accurate depiction of the data. It seems the authors may have done this for the sake of keeping the math and code simpler, which is fine since the assumption is clearly stated, but then it seems unfair to imply that a similar approach by other modelers is

problematic (lines 51-52)."
Response: Thank you for the compliment on the clarity of the descriptions provided. We are sorry if some of the some of the assumptions made in the manuscript came across as unjustified. We will better highlight the assumptions to users, specifically which assumptions apply to the developed method in general (lines 91 f), and which ones were made specifically for the examples (lines 392, 459 – 460).
Changes made: Clarified what assumptions hold in general (line 94, lines 120 f) and which assumption are introduced for the sake of the example (line 415). The mention of "data driven rather than assumption driven" was removed (line 97 and 697) and replaced by a discussion about uncertainty partitioning (lines 94 ff, 426 ff)

Comment: "I would also suggest that the authors be careful about the wording they use to describe their methods in relation to other methods. I suggest moderating some of the language throughout the manuscript to get away from calling other work a "crude simplification" (line 6) and other terms that can read as pejoratives. [...]"
Response: As mentioned in the responses to Robin Trayler and Maarten Blaauw, we will adjust the phrasing in the revised version of the manuscript and apologize if this came across as pejorative.
Changes made: The language was adjusted throughout the manuscript (e.g., lines 3, 4)

Comment: "One thing that could strengthen this manuscript is the addition of a paragraph or two within the introduction that talks in more detail about what we know about sedimentation rates. What kind of distributions are observed? What kind of variability in rates has been documented? The authors state that "it is not clear that in all such cases the assumptions of these age-depth modeling procedures are suitable representations" (lines 49-50). Could they expand on this? What, specifically, do they know that suggests these assumptions are incorrect or limiting? Can they reference particular empirical studies about sedimentation rates? I do not have a particular reference in mind; I just would guess that there exist studies that use physical or computational models or modern depositional settings to measure sedimentation rates. This could help justify the need for FAM and ICON and provide context that helps the reader assess whether common assumptions about sedimentation are so simplistic that they significantly limit the accuracy of age-depth models. In general, working to not only state assumptions but justify them would help strengthen this manuscript."
Response: Based on your and the other reviewers comments we agree that this point was not sufficiently explained. We will expand the text about sedimentation rates in the introduction.
Changes made: Expanded the introduction with more context on what we know about sedimentation rates and their change across environments and timescales (lines 74 ff. 49 ff.). This is still an area of ongoing research (lines 606 ff.)

Comment: "I think the worked examples (beginning of section 3) are a particularly useful part of the manuscript. These examples are likely to be of interest to the reader, not just those interested in the Late Devonian and PETM, but all who want to see the benefits of applying these methods. Because of that, I wished these worked examples appeared earlier in the manuscript to capture the interest of the reader who wants to see that applicability.
Relatedly, I find the math easy to follow, but I think that the math-heavy parts of the text in Section 2 could be moved to the supplementary materials. While the reader may be able to follow the math, the anticipated user the authors specified in lines 81-83 likely won't need the derived equations to implement FAM and ICON, and they would likely be willing to access the supplementary materials if they needed that information. However, I do think there are parts of

Section 2 that can be preserved in the manuscript without negatively interrupting the flow for the reader. For example, I found lines 165-180 of Section 2.4 useful for connecting this work to previous work and in seeing how FAM fills a gap."

Response: As mentioned in the response to Jarno Huygh, we believe that thorough documentation of mathematical structures underlying methods are important and should be an integral part of a methods manuscript. We relay this decision to the editor and are prepared to move the mathematical derivation into the supplementary materials if they believe this would improve the readability, clarity, and absorption of the manuscript.

Changes made: We have left the decision to move the mathematical part to the supplementary material to the editor, who did not make this request. We have decided to keep all mathematical derivations in the main text body, as we maintain that methodology should be thoroughly documented.

Comment: line 25: "The manuscript states "Although age-depth models are rarely the focus of standalone publications" and then later cites many age-depth model papers (e.g., lines 36-38) and other papers that use age-depth models to analyze a sedimentary record or time period. I wasn't sure what the purpose of this statement was."

Response: The point made is that although age-depth models are crucial for temporal inferences, they are rarely published as a standalone result. We will adjust the phrasing and the cited publications to strengthen this point.

Changes made: Removed the first part of the mentioned sentence to streamline the point made (line 28).

Comment: Line 165: "Consider "It is common to assume that..." instead of "Assume that..." With how this sentence is currently written, when I started reading the paragraph, I was not sure if this assumption was a common practice or some sort of thought exercise."

Response: This paragraph introduces the general principle underlying FAM using the conceptually simple example of constant tracer flux, which is relaxed later. The "Assume that" introduces the example case of constant flux. We will rephrase the paragraph to clarify this.

Changes made: Changed to "given" to clarify that the tracer fluxes are needed as input for the method (line 194, 216).

Comment: "Line 334: Define FAIR as Findable, Accessible, Interoperable and Reusable within the parentheses"

Response: We will add this in the revised version of the manuscript.

Changes made: Adjusted (line 371)

Comment: "I assume many of the citation-related typos will be caught during copy editing, but I wanted to point towards a couple: Lines 415 and 653: Gradstein (2020) should be Gradstein et al. (2020), Line 496: It seems like Ramsey (2009) might be missing from the list"

Response: Thank you for catching those typos, we will correct them in the revised version of the manuscript.

Changes made: The typos were corrected.

Comment: "Some sections seem like they might need a renumbering to maintain a logical hierarchy to the sections. For example, the Late-Devonian and PETM examples should probably be at the same level within the numbering scheme. Line 437: Consider renumbering Section 3.1.2 to 3.2, Line 487: Consider starting a new section (Section 4) for the Discussion since it discusses more than just the worked examples"

Response: We will adjust the hierarchical ordering of the sections in the revised version of the manuscript

Changes made: Hierarchy of sections was adjusted throughout the manuscript

Comments by Robin Trayler

Comment: *"This manuscript presents some interesting methods for constructing age depth models, and the underlying code appears to be very flexible and allows users to specify arbitrary distributions and uncertainty structures. I view this as both a positive and negative. As is, it is difficulty to tell exactly how one might actually generate an age depth model and what information is needed. For example, the ICON model relies on sedimentation rate estimates, but it is unclear where these come from and what they are based on. Given that one of the most common uses of age-depth modeling is to estimate sedimentation rate, it is unclear what a user is gaining from using these methods over more established models."*

Response: There are many types of information on the distribution of time in a section that come from sedimentological and stratigraphic sources. They include, but are not limited to, estimates on sedimentation rates, e.g., from evolutive procedures in cyclostratigraphy such as eCoCo (Li et al, 2018) and eTimeOpt (Meyers 2019), relative changes in sedimentation rate based on sequence stratigraphic information, typical sedimentation rates associated with facies, or dilution or condensation of tracers (Jarochowska et al. 2020). We believe this type of information is currently underused in age-depth modelling. Our methods allow to formalize this knowledge and turn it into quantitative age-depth models. Specifically for cyclostratigraphy, where uncertainties from evolutive methods were previously not propagated into age-depth models.

We will be more explicit about the potential sources of information that can be used to construct age-depth models via the proposed methods, and improve signposting to external materials associated with the package (e.g., vignettes) that expand on the computational details.

Changes made: Expanded the introduction to include more references to sources of information that can be turned into age-depth models with the described methods (lines 74 ff 75). References to the latest version of admtools were added, which includes expanded documentation and vignettes.

Comment: *"I also agree with Dr. Blaauw's comment that the manuscript "[has] an axe to grind with regard to existing approaches to age-modeling" which I think harms the message of the manuscript. I feel like the manuscript would be well served to focus on comparing and contrasting their methods with existing methods, rather than dismissing the existing methods out of hand as "crude simplifications"."*

Response: We apologize for how this came across and realize this was inappropriate. We will adjust this in the revised version.

Changes made: Adjusted the language throughout the manuscript (e.g., lines 3 f)

Comment: *This manuscript makes some strong statements about "assumptions" throughout that I find problematic. especially since the methods presented within essentially contain equivalent assumptions about sediment accumulation. I appreciate that the manuscript lays out their basic assumptions about sedimentation in section 2.1, but this does not mean they are not being made nor does it inherently make these methods more data driven. It only means that* *they are being stated clearly, as they have been the papers detailing past age-depth models [@haslett2008; @blaauw2011]. For example, I view assumption 1 as a valid compromise for constructing age depth models, but it is not strictly true in a geologic sense. Hiatuses and negative accumulation (erosion) are both clearly observable processes but are difficulty to deal with from a modeling standpoint. Nevertheless, multiple existing models (Bacon, my own astroBayes) have at least rudimentary ways of dealing with these problems.*

Response: Our intention was to convey that we do not assume any **processes or distributions** on sediment accumulation, which is the difference to the models we compare with. We did not claim not to make any assumptions at all and that seems to be a misunderstanding, showing that we need to make it more clear. Our assumptions are the minimum subset of assumptions made by other models, but without those on the distributions of gaps or distributions from which sedimentation rates need to be drawn (please see the discussion with reviewer Maarten Blaauw). In other words, we strip the software from embedded assumptions on parametric processes and create the **minimum** model that would still estimate the ADM, and to which further assumptions can be added in a modular way. These assumptions may be added by the user and the trade-off between the number and confidence in assumptions and the precision of the ADM can be evaluated.

We realize a distinction should be made in the text between the FAM and ICON models and the software itself. Assumptions listed in 2.1 are of the models, but admtools handles hiatuses. So the ADM inference using either of the models can be done in a piecewise manner.

Changes made: The role of assumptions in the model development was clarified in the introduction (lines 47 ff, 63 ff)

Comment: *"The manuscript does a very thorough job laying out their mathematics for both the ICON and FAM models. However found much of this information difficult to follow. As written, it's difficult to tell what information is needed to actually construct a model. Is it a series of dates with stratigraphic positions? Continuous or discrete estimations of sedimentation rate or tracer flux? I would strongly suggest the inclusion of a few conceptual figures that show how the models are actually constructed in practice."*

Response: "Thank you for your comment on the thoroughness. We believe detailed description of scientific procedure is crucial for reproducibility and to open up black boxes. We will be happy to include some conceptual figures into the revised manuscript and provide more detailed instructions on what information is needed for the methods. As is, most of this information is associated with the package itself as vignettes (long form documentation), as we believe that this is more user-friendly. We welcome any feedback on these materials via GitHub."

Changes made: None

Comment: *"The manuscript follows best practices for code and data availability. The linked Zenodo repository contains a fixed version of the R package and the CRAN/ Github version are also well documented."*

Response: Thank you. We value thorough documentation scientific methods.

Changes made: None

Comment:*"I would like to see more details on how the models were tested however. First how has it been verified that equivalent results are generated each time the models are run? As far as I can tell, there are stochastic, random number generators "under-the_hood". Do these always produce the same answer (allowing for small amounts of change due to the nature of random number generation)?"*

Response: As with all Monte-Carlo methods, small random fluctuations are expected in the results. For full reproducibility, users should fix their seed for random number generation (as is standard practice) and use computational environments via the renv package to minimize computational differences arising from different R or package versions (a practice that is standard in Python, but not yet very common in R). Because these are all standard practices in scientific computing, they were not specifically mentioned in the manuscript. We will add these practices to the documentation for users not familiar with them. The unit tests can be inspected in the R package in the test folder. They test edge cases where analytical solutions are known (e.g., constant sedimentation rate).
Changes made: None, basic principles of computational reproducibility are beyond the scope of this paper.

Comment: *"Second, how do these new methods compare to established age-depth models? For example the Bacon and Bchron R packages, and OxCal have 10,000+ of citations between them and have become the de facto standards for age-depth modeling in many contexts. How well does admtools compare to these packages? Does it produce more precise models, allow fitting of information not allowed by them, etc? I'd like to see this explored more systematically to help guide users towards the most appropriate models for their data. It might be worth doing comparisons of admtools with Bacon/Bchron/ OxCal for the Devonian or PETM examples. This could be added to the "comparison with other methods" section"*
Response: As you mention, the main difference between existing approaches and ICON is that ICON uses existing information on sedimentation rates to construct age-depth models as opposed to estimating sedimentation rates (your comment: "For example, the ICON model relies on sedimentation rate estimates, but it is unclear where these come from and what they are based on. Given that one of the most common uses of age-depth modeling is to estimate sedimentation rate, it is unclear what a user is gaining from using these methods over more established models."). For potential sources of such information see our response above. We will clarify this difference in the "comparison with other methods" section, and highlight the scenarios where the presented methods are of use.
We argue that age-depth models are hypotheses on sediment accumulation, and should as such not evaluated by their uncertainty, but in light of our understanding of the depositional environment of interest (section 3.2.3, lines 543 onwards). The presented methods construct age-depth models from available information on these depositional systems. As such, discrepancies in age-depth models between our approach and previous methods (e.g., Bchron) would highlight discrepancies between the sedimentation models used by these approaches and our understanding of depositional systems. While we believe such a test of model adequacy would be beyond the scope of this paper, we think that it is an interesting question in and of itself and agree to expand the manuscript accordingly if the editor requests it.
Changes made: Because the editor has not made the request to add a comparison of different models and a discussion of model adequacy, none was added. The point that this topic requires more research was added to the discussion (lines 606 ff., 664 ff.)

Comment: *"Line 375: There is a probability based version of timeOpt included in astrochron. I am unsure if it has been officially implemented in eTimeOpt however. See [Meyers and Malinverno, 2018] for details."*
Response: Unfortunately timeOptMCMC has not yet been implemented as an evolutive procedure. We believe it would solve the problem mentioned in line 374 (a probability density for sedimentation rates). As is, timeOptMCMC provides only a posterior distribution of the

sedimentation rate for a whole section.
Changes made: None

Comment: *"Figure 2: I think the floating and anchored subplots are labeled backwards or It is not clear how the anchoring date is being used. I would suggest double checking the labeling and adding a point + error bars to the anchored panel to visually show the anchoring date and its uncertainty. As is, it's very unclear why the "floating" model would converge to near zero uncertainty in the bentonite layer and the anchored one would not. Is this convergence because the "floating model" is actually assuming the bentonite has a fixed age with no uncertainty?"*
Response: We are sorry if the terminology is confusing, and will expand this explanation in the revised manuscript. The age-depth model in A is "floating" as it is not anchored in absolute age, but time is measured relative to a fixed stratigraphic tie point (the bentonite layer). The age-depth model in B is anchored in absolute time via the absolute ages from Percival et al (2018). This corresponds to the usage of the term as used in Hinnov (2013). It is indeed true that the uncertainty in the floating age-depth model is 0 at the bentonite layer, as time is measured relative to it (shown by the negative times below it, indicating time before the deposition of the bentonite).
Changes made: The floating age-depth model was removed and replaced by one anchored at the mean U-Pb date. The text was adjusted to better highlight partitioning of uncertainty derived of different sources (lines 426 ff., Figure 2).

Comment: *"The bentonite uncertainty (±0.053 Ma; 1 or 2 σ?) should have error bars that cover slightly less than the distance between the a pair of grid-lines on the x-axis."*
Response: The uncertainties given here are 2 sigma. We will add the error bars to the plot.
Changes made: The plot was modified to include the 2 sigma uncertainty and show this uncertainty propagating into the section (Figure 2)

Comment: *"Line 395: how can the model be "floating" if it has a tie point?"*
Response: The age-depth model is floating in the sense that the absolute age is unkown, but age is measured relative to a fixed stratigraphic tie point (here the bentonite layer). See also our comment above.
Changes made: The floating age-depth model was removed in favor of one anchored at the mean age of the radiometric tie point. The text was expanded to highlight the original point we wanted to make, which is the partitioning of different uncertainties and their propagation into the section (lines 426 ff, Figure 2)

Comment: *"Line 519: A minor comment but it cannot incorporate gaps in the stratigraphic record. The presence of fossils in the rocks seems to play no role in model construction."*
Response: Thank you for catching this, we will correct this in the resubmission.
Changes made: Corrected

Comment: I have one minor comment regarding my own work. On like 500 of the preprint the authors state:
"For example, the computations accompanying Trayler et al. (2023) publication of astroBayes will take "several days or weeks" to compute on a laptop. In contrast, the computation time for the examples show here is typically within minutes.".
The "days or weeks" quote does not come from the main manuscript, and is instead found in the github repository (https://github.com/robintrayler/astroBayes_manuscript) associated with

model testing and validation. This code generates ~10,000 individual models (each made up of thousands of MCMC iterations) and was intended to test and validate overall astroBayes model performance. *Individual* astroBayes models take a few minutes to maybe an hour, depending on complexity and the number of parameters. It is somewhat disingenuous to compare the time it takes to compute thousands of models to the time it takes to generate single models using the authors code.
These computational times are also true for the other Bayesian models mentioned by the authors. Bchron, Bacon, OxCal, etc can all be run on a personal laptop in a few minutes to hours depending on complexity and the number of user specified MCMC iterations.
Response: We apologize for this mistake, and will remove the corresponding part when revising the manuscript.
Changes made: Removed the discussed part (lines 590f).

Comments by Maarten Blaauw

Comment: *„I wonder though how many researchers will have the expertise to select the parameters and settings appropriately/correctly. Is there a danger that users will (inadvertedly) choose settings such that they get age-models based on wishful-thinking?"*
Response: We advocate a multiple hypothesis approach in age-depth modeling. In this case: estimating the age-depth models under different assumptions. Given the uncertainty of the assumptions on the distribution of sedimentary and reworking events in the geological record, a sensitivity analysis would be advisable in each case where there is risk of diverging conclusions under different ADMs. This was the intention behind delegating the decision to the users: unlike in packages, where assumptions are embedded, in admtools users can (and are encouraged to) assess their impact.
Changes made: The underlying philosophy of the approach was highlighted in the manuscript: Users make their assumptions explicit and can build modular models to study the relative contribution of uncertainty on the age-depth model (lines 64 ff, 94 ff 659 ff)

Comment: "The usage in the abstract of terms like 'simplistic' and 'crude simplification' should be toned down."
Response: We apologize for how this came across and realize this was inappropriate. The wording will be changed in a revised version of the manuscript.
Changes made: The language was adjusted throughout the manuscript (e.g., lines 3, 4).

Comment: "It is frustrating that the manuscript rejects these mathematical distributions but then proposes a uniform distribution for sediment accumulation rates, which to me seems much less geologically robust/defendable and almost entirely made for ease of coding."
Response: We apologize if our comment came across as rejection of the use of gamma distribution for sedimentation rates: This will be rephrased in the revision. We acknowledge that the gamma distribution has been well investigated and carefully chosen for a particular type of environment and timescale (in particular, peat bog cores). Our concern with the gamma distribution solely pertains to its use „across the board": For all different timescales and for all possible different sedimentary environments (e.g., pre-Quaternary sedimentary sequences). We are standing on the shoulders of the empirical work undertaken by the referee and by Goring et al. (2012), and we will make this acknowledgement more clear in the revised version of the manuscript. We solely want to point out that such empirical studies have not been repeated for a wide range of depositional environment and timescales.

Some depositional environments are characterized by rapidly changing sedimentation rates. In such cases, estimating the distribution from which sedimentation rates of depositional events are derived is challenging: It is only possible in cases for which the discretization of depositional events is possible (e.g., bedding) and the position and duration of hiatuses are either negligible or very well constrained. In all other cases, it is not possible to ascertain the distribution from which sedimentation rate is drawn, nor its parameters. We strongly support the need for more research on informing the parameters used to estimate age-depth models across depositional environments.

The goal of admtools is to allow all options, with the uniform distribution being arguably the least obliging. Based on the reviewer's recommendation, a gamma distribution utility will be included in the next admtools release, which will be prepared alongside the revised manuscript. Nevertheless, we insist on distributing it with the caveat explained above. We want to highlight that the admtools package allows users to specify any uncertainty on sedimentation rates, be it gamma, uniform, or any other (potentially nonparametric) distribution. As such, the utility function to use the gamma distribution is a special case of the more general functionality provided.

Changes made: The gamma distribution was implemented, which is now mentioned in the manuscript (line 355). The role of different depositional environments was highlighted in the introduction (lines 47ff), the abstract (lines 7) and the discussion (lines 603)

Comment: "Lines 403-6, limits on sedimentation rates: how are these decided? For your chosen limits of 0.1 to 0.6 cm/kyr, is it really geologically likely that any value inbetween that range is equally likely, but that any value outside of that range has a probability of 0%? I don't think so."

Response: The bounds follow the original study by Da Silva et al. (2020), we used them to keep the example close to the original publication. But please note that the sedimentation rates are not modeled here to be uniform between these limits, but the frequency distribution of sedimentation rates (piecewise for the windows used by TimeOpt) is derived from eTimeOpt based on its (the astrochron package's) r^2_{opt} metric. We recognize that this needs to be made clear in the text.

Changes made: This was clarified (lines 447 to 455)

Comment: "Why not use say a gamma distribution which by itself enforces sedimentation to never be <0, and which can downweigh sedimentation rates that are considered to be unlikely, without the need to enforce such strict borders? With more information available, the gamma distribution can be made to peak more (or less)."

Response: The answer in this particular context is provided above: here the distributions of sedimentation rates are extracted from eTimeOpt and thus the assumptions are delegated to that method.

Changes made: None

Comment: "Line 70, all models are partly driven by data, and partly by assumptions."

Response: We agree that the sentence in line 70 is unnecessary.

Changes made: This was removed (lines 97 and 697)

Comment: "Lines 531-536, yes, ages often scatter beyond their lab uncertainties, but Bayesian models deal very well with such scatter - much better than classical models can. Same for lines 551-551: both are examples of classical models, yet real-world data and simulations have shown that Bayesian models provide much more pessimistic/realistic uncertainty estimates (e.g., Blaauw et al. 2018 <doi:10.1016/j.quascirev.2018.03.032>). How do the uncertainties of

the proposed age-depth model compare with existing Bayesian/classical models?"
Response: In the revised version, we commit to making clear that admtools is not a criticism of Bayesian approaches to age-depth modeling. For many situations where the depositional environment and sedimentation dynamics offer informative priors, these approaches are supreme. Our work is mainly inspired by applications to deep-time and long intervals, where such priors are not available. We concur that in cases when Bayesian methods are applicable, they should be applied. As mentioned earlier in this answer, further collection of parameters of sedimentation dynamics across environments would hugely improve the assumptions of either method. The point made about mixing in lines 531 to 536 is not a criticism about any specific methodology for age-depth modeling. The limitations of age-resolution imposed by mixing are well-studied, and show that any sedimentary layer has no unique age (as assumed by most age-depth models), but rather a range of ages determined by the physical processes of sediment deposition and mixing. This is an important limiting factor for age-depth modeling, which is why we discuss it here.
The examples in lines 551 serve to demonstrate that there are logical and intuitive constraints on uncertainty of age-depth models, a point that holds independent of the methodology used to estimate them. We argue that age-depth models are hypotheses on sediment accumulation history, and should as such be evaluated in light of our understanding of the corresponding depositional system, not by means of its uncertainty alone.
Changes made: Clarified that Bayesian age-depth modeling tools will automatically resolve reversals, but the depositional resolution of a given environment will be a limiting factor on the maximum resolution achievable in an age-depth model.

Comment: "Line 54, note that orbital matching (as opposed to age-depth modelling based on e.g. radiometric dates) makes assumptions that could be seen as simplistic, hard to prove, and causing a degree of circularity with and dependence on other records. See Blaauw 2012 <doi:10.1016/j.quascirev.2010.11.012> for a critical review of climatic tuning - although I agree that sometimes one would need to be pragmatic and choose tuning if no absolute age estimates are available. "
Response: Our example is not intended to argue that "orbital matching" is better than any other method; the assumptions and limitations of that approach are outside of the scope of admtools and are well argued in the cited paper. We favor treating age-depth models as hypotheses and entertaining different tools to assess the uncertainty of the inference. Models derived from astrochronology should be compared with alternative approaches and with models derived for slightly different parameters, when possible, to assess the sensitivity of the model. In this example, we refer to the original publication of Da Silva et al. (2020) for the justification of this methodology; the terms of its use still apply.
Changes made: Added a word of caution about orbital matching to the introduction (lines 73 ff)

Comment: "The CON algorithm is presented using many equations, but in the end defines nothing fundamentally new - this has all been stated before in different notation or terms in previous age-depth model papers."
Response: The reviewer is right in stating that the equations bring nothing fundamentally new to the table. These equations are scattered over the literature in different forms. We prefer to provide full disclosure of the mathematical treatment of the input ages, depths, and sedimentation rate constraints in the ICON algorithm. In other words, we provide a general purpose implementation to construct age-depth model from arbitrary information on sedimentation rates throughout a section, opening the opportunity to turn previously unused stratigraphic information (e.g., semiquantitative changes in sedimentation rate from sequence

stratigraphy, estimates on sedimentation rates from cyclostratigraphic methods such as eTimeOpt or eCoCo) into age-depth models, including propagation of all involved uncertainties into the age-depth estimates. The high level of mathematical formalism in this is to ensure the method is thoroughly documented in its generality.
Changes made: None

Comment: "As for FAM, another classical use of correcting for supposed constant influx assumed a constant pollen rain and then applied this assumption to stretch and compress a core's accumulation rate. See Middeldorp 1982 <doi:10.1016/0034-6667(82)90003-3> and Young et al. 1999 <doi:10.1016/S0034-6667(98)00060-8>. Of course, on long time-scales it is unlikely that any proxy will have a constant flux (Figure 4 and the accompanying discussion are interesting in this respect). The interpretation that the authors' FAM model essentially resembles the CRS model doesn't make the new model novel either - the CRS stems from the 1970s. "
Response: The affinity with the CRS model is stated in the text, but here it is generalized to different types of fluxes. Constant is used in the example, again following a published study, but a different relationship – such as exponential decay – can be input by the user. We agree with the referee that these references should be added. As stated in lines 165 to 180, we are well aware that this approach has been used before, and we provide a general description and implementation of the procedure without constraints on the tracer fluxes. The constraint on constant tracer fluxes has always been a criticism of this approach, as pointed out in the example regarding the PETM. Among other things, our approach allows to assess the robustness of this assumption (see PETM example), incorporating external knowledge on tracer fluxes in the procedures (e.g., based on expert knowledge), making it an extension of existing methods.
Changes made: Reference to the mentioned publications was added (lines 197 ff, 675 ff)

Comment: "Line 301, "Timing and positions of tie points can follow arbitrary probability distributions as long as they are strictly ordered." Do you mean that closely spaced dates/tie-points cannot have overlapping age distributions? That would be a drawback of the method. This type of problems is exactly why Bayesian models which use Poisson/Gamma distributions are so successful in modelling sedimentation - they guarantee that reversals are absent, and moreover, they can take advantage of high dating densities; they "learn" with precision becoming ever higher as more dates are added. This is not necessarily the case for classical age-models (see Blaauw et al. 2018 <doi:10.1016/j.quascirev.2018.03.032>). "
Response: The answer to the question "Do you mean that closely spaced dates/tie-points cannot have overlapping age distributions?" is: no, it does not mean that. In the presented methods, the user has to decide how to resolve age-reversals that arise from overlapping age distributions. One option implemented in the provided templates is to discard non-monotonous samples drawn from those distributions. true age-reversals of multiple decimeters are well documented in Holocene marine environments based on AAR dating of individual shells (e.g., Kosnik et al., 2007; Kosnik et al., 2009; Dominguez et al. 2016, Tomašových et al. 2018). Admtools gives the users the option to make an active choice how to handle such scenarios.
Changes made: A sentence on the physical limits of age-depth inference and their relation to depositional resolution was added (lines 586 ff.)

Comment: "Fig. 2, isn't the floating age-model in panel B and the anchored one in panel A? If not, then the terminology used is confusing (at least to me). If a core's age is measured relative to an otherwise dated layer (i.e., it is tied), how can it be floating? Surely the age uncertainty of the Bentonite layer is not 0, so there is no point in presenting Figure 2A. (Later on in the Discussion, a point is made that one can set the uncertainty at a layer to 0 in order to estimate the durations of an interval. However, that can also be done by simply subtracting the bottom ages from the top ages for each iteration of say a Bayesian age-depth model, be it floating or anchored.) For Figure 2A, I also don't see a 'sausage-shape' but rather a nematode-shape. Are the reconstructed 95% ranges of Fig. 2B considered realistic? Do they compare well with established methods such as BChron? This should be discussed, e.g. after line 511."

Response: We are sorry if the terminology is confusing, and will expand this explanation in the revised manuscript. The age-depth model in A is "floating" as it is not anchored in absolute age, but time is measured relative to a fixed stratigraphic tie point (the Bentonite layer). The age-depth model in B is anchored in absolute time via the absolute ages from Percival et al (2018). This corresponds to the usage of the term as used in Hinnov (2013). The point of figure A was to show that even in the absence of absolute age-constraints, information derived from sedimentation rate estimates can be used to constrain durations of intervals, and are not limited to the uncertainty of the absolute ages used (lines 425 following).

Regarding the comparison with BChron, see the comments above – ICON uses information on sedimentation rates from sedimentological and stratigraphic sources to construct age-depth models, whereas BChron estimates sedimentation rates.

Changes made: The floating age-depth model was removed and replaced with one anchored at the mean age of the U-Pb date. The purpose of this age-depth model (partitioning of uncertainties) was expanded upon (lines 426 ff, Figure 2). Regarding model adequacy: We left the choice to discuss model adequacy and comparisons with age-depth inference methods to the editor, who did not request this change. A comment on future research required in this direction was added (lines 606 ff.)