The reviewers comments for our original manuscript were extensive and helpful. However, they were often conflicting with the 6 different reviewers focusing on different aspects of the manuscript. The comments were all very helpful and have significantly improved the manuscript. We have addressed these comments in the revised manuscript. Here we provide a summary of the reviewers' comments and then detailed points below in alphabetical order.

Alexis Ault provided helpful comments and highlighted where the manuscript was confusing. She misunderstood a key part of the paper and did not appreciate that all the diffusion data was used while fitting the data. We have clarified this point to make it more clear that all the data are used to constrain the uncertainties in the model.

Rebecca Flowers helped clarify lots of the language to ensure that the background information was correct. This has greatly improved the manuscript.

Willy Guenthner pushed us to incorporate the uncertainties directly into QTQt and this was a major undertaking. These new models form a new section of the paper and a modified version of QTQt is now available from Kerry Gallagher.

Kip Hodges highlighted other relevant research but also helped us to frame the community recommendations more clearly. It is vital to carry out additional diffusion experiments and continue to refine the models, but in many cases the models appear to do a good job at matching geological constraints despite uncertainties.

Brenhin Keller's review was very helpful and questioned our use of QTQt. We ran longer models and this did have a small impact but we have modified our QTQt models so that they now predict a GU. He also helped us clarify some of the subtle concepts debated when considering the origin of the GU.

Kalin McDannell's review was also very helpful and appreciative of the need to assess uncertainties in the model. We have rerun the QTQt models to more closely match the published models.

Detailed responses to these reviews can be seen below.

Many thanks,

Matthew Fox

Reviewer: Alexis Ault

Our responses are to the reviewer's comments are shown in italics.

The manuscript "Origin of Great Unconformity obscured by thermochronometric uncertainty" by Fox et al. presents numerical models of uncertainty of He diffusion kinetic parameters from an undamaged zircon and zircon N17 (of Guenthner et al., 2013). Specifically, they consider uncertainties in the frequency factor (D0) and activation energy (Ea) for these two end-member grains, which co-vary, and they use ZRDAAM (Guenthner) with their updated kinetics for these two grains to predict zircon (U-Th)/He (zircon He) thermochronometry dates for an example deep time thermal (tT) history.

The reviewer comments are very helpful and highlight ways that our paper is unclear. However, several of the comments the reviewer makes are not correct and this is probably due to a lack of clarity on our behalf. We apologise for this confusion. In particular, we do not consider uncertainties "in the frequency factor (D0) and activation energy (Ea) for these two end-member grains". These end member grains are the basis of the entire radiation damage model calibration and we discuss the implications for this methodological misunderstanding below.

This contribution drives at a timely and important topic – the utility, and potential limitations, of zircon He thermochronometry to resolve deep time thermal histories. As the authors point out, zircon He data has been used to evaluate the timing and magnitude of erosion during the creation of the Great Unconformity, with data patterns being used to discriminate between glacial, tectonic, and geodynamic mechanisms for unroofing. However, uncertainty in He diffusion kinetic parameters, which are a function of radiation damage accumulation and annealing, have not previously been considered before and I appreciate the authors tackling this here. As the authors show with their suite of numerical models: the effect of these uncertainties is magnified in long time scales (i.e., deep time) and dispersion in a zircon He date for a given eU value and deep time tT history can be >>> 100 Ma. This suggests that, in the absence of additional diffusion experiments on zircon grains over a spectrum of effective damage, it may not be possible to resolve cooling during say the Sturtian and Marinoan glaciations, or doing a specific Neoproterozoic time period.

I think that this manuscript is suitable for Geochronology pending moderate revisions. There are elements of the manuscript that I think should and can be expanded and clarified prior to acceptance for publication:

First, the authors center their analysis on only two end-member zircon grains. One grain, "z", is sometimes called a minimally damaged grain and elsewhere called an undamaged grain or a "low" grain. As Guenthner et al. (2013) shows, these are different. What is the source of the data for the low/undamaged grain? This is never stated in the text and should be added. (see comment below about Figure 1).

The data used for the analysis are the diffusion data used in Guenthner et al. 2013. We use all the same data and analyse all the diffusion data from all the crystals *simultaneously. The undamaged grain does not exist and its diffusion kinetics are determined by extrapolating the damage model. This is the same as is done in the original Guenthner paper.*

Second, and more importantly, it is not clear how the authors are taking accounting for uncertainties in intermediate or even high damage zircon grains (N17 is effectively an amorphous grain; Guenthner et al., 2013). The authors state (152-153) that they are focused on end member grains and imply their approach is an improvement on the extrapolations between high and low damage levels in Guenthner et al. (2013) (152-157). Regardless of whether the low damage end member is zero damage or Mud Tank kinetics, what happens in between the low damage and N17 is, I think, critical. What about moderately-damaged to high damage grains: how do the predicted versus observed diffusivities vary for RB140, BR231, M127, and G3? (Guenthner et al., 2013)? Note that Guenthner et al. (2013) presents data for slabs of the crystals cut parallel and orthogonal to the C axis for each grain. Related to this, looking at figure 3, I found myself wondering if D0 and Ea are correlated in the same way in moderate to high damage zircon grains, and do these correlations align with Guenthner et al. (2013), as they do for the low damage "z" grain?

All of these data are used in the analysis. Our approach is designed to maximise information from the correlations that were neglected in the original Guenthner et al., (2013) study. We also show the fit of the model to the data.

The reason why this additional analysis I think matters is because the evolution of effective damage and resulting He diffusivity is complex as the authors point out (126-128). Diffusivity initially decreases from an undamaged/low damage grain to a moderate damage grain, but diffusivity then dramatically increases from a moderate to an amorphous grain. In figure 4, the greatest spread in dates when taking into account the variation in grain "z" and N17 uncertainties occurs in the range of moderate to high damage grains (1000-1500 ppm eU). If only endmembers are being considered, how are uncertainties in diffusion kinetics between these two endmembers being accounted for or propagated through a tT history? This was not clear to me.

All the data are used to constrain the damage model and we have attempted to highlight this point in the revised manuscript.

I think the authors can address these questions in two ways: I encourage the authors to perform the same calculations that they perform on an undamaged/low damage grain and N17 on additional grains, mentioned in the paragraph above, and expand figure 2. I think showing how the D0 and Ea values compare with Guenthner et al. (2013) will be useful for the community. It was interesting to see how their models reproduced N17, but not for the un/low damage grain. At the very least, the authors should expand the text to explain how they are interpolating between un/low damage and the amorphous grain when propagating model uncertainties across a range of eU – I apologize, I could not follow how this was done with the information provided.

We have provided more information in the paper to help explain this analysis. We have highlighted that all the crystals are used in the likelihood function and described how the damage model itself is used to interpolate between the values.

Third, Figure 1 would benefit from more information. Adding these details will help the reader follow the discussion in the text and also help if additional analyses are added:

-Please label each of the Arrhenius relationships for the different diffusion experiments – perhaps using different symbols and colors. I think these figures could be enlarged as well. I found myself flipping back to the original data of Guenthner et al. (2013) to identify N17 and I could not figure out the low or undamaged grain.

We have added more detail here and coloured the samples based on alpha dose. We hope that the improved figure helps with identifying grains.

-I would make the trend lines a darker grey or perhaps thicker – it is hard to see them. Enlarging the figures (perhaps single column?) would help.

Additional edits and comments:

12-13: I would rephrase this topic sentence, as it is confusing as written. Perhaps something like "Thermochronology provides a unique perspective on the timing and magnitude of erosion during the generation of unconformities."

Corrected.

17: I do not think However is needed at the beginning of this sentence.

Corrected.

23: I would add "… as a function of radiation damage accumulation and annealing." At the end of this sentence.

Corrected.

25: "less well resolved" – during what time interval?

By not specifying a time interval, we are highlighting that the whole history is less well resolved.

26 and 320-322: "calibration with natural laboratories" – Calibrating zircon He diffusion kinetics from natural samples that experienced deep time thermal histories is challenging, right? Because it is precisely during those long tT histories where issues such as U-Th zonation will be exacerbated. And so if you chose to advocate for an empirical approach, I would add a cautionary note that provided U-Th zonation can be quantified in analyzed grains (which is hard, right? because grains that are consumed for zircon He analyses are not available for detailed LA-ICPMS zonation investigation.

And LA-ICPMS zircon He analyses also have uncertainties that would need to be taken into account in this modeling approach).

33: I would omit Keller et al., 2019, as this Great Unconformity study does not include zircon He thermochronometry data, which is the subject of the sentence.

Most of the introduction has been rewritten and shifted around. We still cite this paper because it is an important contribution to the entire debate.

36: I believe this is 1300 to 200 million years of missing time.

Thank you!

38: Here I think you could cite other papers besides McDannell et al., 2022 such as those that you site in 33 plus others, minus Keller et al., 2019.

Corrected.

38-40: Although I do not think you want to wade into this with this paper, I would rephrase this because the modeling approaches described in McDannell et al. (2022) reflect only one approach to this problem as I understand it.

This section has been rewritten.

95: “variability in radiation damage annealing parameters” – do you mean “diffusion kinetic parameters”? Although D0 and Ea depend on the effective radiation damage (i.e., accumulated and annealed damage), the work presented here does not evaluate damage annealing specifically as this phrase implies.

Corrected.

210: Useful to call-up the inset in Figure 4.

Figure 4: Please add tick marks for x and y axis values.

229: Grain size = equivalent spherical radius? Or?

Corrected on line 167.

244-245: These are incredibly precise numbers – is this level of precision warranted here? And how much does it matter?

We have simplified these numbers, but we do want to be accurate and have displayed the full distributions.

238-256: I think this analysis would be useful for grains like RB140, BR231, M127, or G3!

They are done for all analyses.

Alexis K. Ault

Department of Geosciences

Utah State University

Reviewer: Rebecca Flowers

Our responses are to the reviewer's comments are shown in italics.

Review of Fox et al., Origin of Great Unconformity Obscured by Thermochronometric Uncertainty, Geochronology

We thank Prof. Flowers for this detailed and helpful review. We have modified the text where suggested and this has improved the framing of the problem and also the ongoing debate. We have attempted to satisfy all reviewers with our modifications. These are listed in the general response we have provided. Due to the number of changes requested by the reviewers, we have not tracked changes.

The core of the discussion focuses on whether the thermochronometry alone can be used to constrain cooling histories through time or whether constraints are required. It is our belief that it is very important to establish what the data can and can not resolve. It is useful to show what a solution might look like without any constraints as in many cases, the constraints are controversial or complicated. Through linking samples in space using concepts like the AER, or grouping samples together, more thermochronometric data is used to infer a solution. This should perhaps be attempted before incorporating additional constraints. It is also important to highlight that incorporating data from different thermochronometric systems will also provide greater resolution.

Our goal was to simply explore the importance of uncertainty in damage parameters. To isolate this, we attempted to infer a thermal history, only using the data. Our model was too simple, however, the best fitting model does approach an acceptable GU thermal history. To ensure that we are able to produce this thermal history, we have decreased the uncertainty on the data. In this way, we have produced a thermal history that is consistent with the GU but have also been able to explore the importance of the uncertainty.

This contribution is focused on better constraining the uncertainty associated with the zircon radiation damage accumulation and annealing kinetic model (ZRDAAM), and then evaluating how this uncertainty influences the ability to resolve thermal histories associated with development of the Great Unconformity. The manuscript begins by reviewing the existing ZRDAAM calibration, presents a new calibration based on the same kinetic dataset as used in the original ZRDAAM, uses a new approach to constrain ZRDAAM uncertainties, carries out inversion modeling of a dataset to explore the influence of ZRDAAM kinetic uncertainties on data interpretations, and concludes with some implications regarding the resolution of (U-Th)/He datasets in deep time.

This is a useful and informative contribution. I agree that there's opportunity to use the current Great Unconformity debate as motivation to improve kinetic model calibrations, which in turn can improve our ability to address problems like the origin of this feature.
This paper seems to approach the Great Unconformity problem from the direction of evaluating whether the thermochronologic data alone can resolve thermal histories in deep time. I completely agree that this absolutely is not possible without integration with other types of information, like that from geologic data.

I provide a variety of suggestions below that I think could be used to further strengthen the paper. Most importantly I highlight (described further below): 1) the need to more clearly and completely explain the new calibration, 2) that it's essential to include the Great Unconformity in any model that is supposed to explain the Great Unconformity, 3) that both the data being modeled and the predictions from the inversion models should be plotted, and 4) one aspect of the implications section that could be more valuable if expanded.
Abstract

Lines 13-14: "…about the origin of the Great Unconformity, a global erosional event that represents a period of almost a billion years at the end of the Precambrian."
Suggest modifying this sentence. The Great Unconformity is a geologic feature, not an event. This sentence also asserts that it represents a global erosional event that represents a billion years – but the debate is actually about the timing and duration and whether it represents a global erosional event or not.
Introduction

Corrected.

Line 33: Keller et al. (2019) isn't appropriate to cite in this sentence because there are no thermochronologic data presented or interpreted in that contribution. Also suggest adding an e.g., before the refs, or adding more refs, because there are more papers that have used thermochronologic datasets to make interpretations about the Great Unconformity than those cited here.

We have modified the introduction and this has been accounted for.

Lines 35-36. "…in the Grand Canyon, the erosional event spans from circa 1200 to 250 million years…". This is the amount of missing time. The erosional event didn't span this entire interval. Also should use an e.g., for the reference, because many have noted this missing magnitude of time in the Grand Canyon.

Corrected.

Lines 38-40. “This approach highlighted that erosion rates increased across the North American craton during Neoproterozoic glaciation, supporting the hypothesis that…”. It would be more appropriate to replace the word “highlighted” with “inferred”. None of the datasets used in that paper require that erosion increased during the Snowball glaciation, as shown in the comment by Flowers et al. (2022). This is true even when one does not take into account kinetic model uncertainties, which I agree further increase the uncertainty in the inferred tT paths, and are the focus of this contribution.

Corrected.

Line 46: “…concepts of geochronology”. Also involves concepts of diffusion. Suggest modifying to “…geochronology and diffusion.”

Corrected.

Line 49: missing a word here, should be “diffusive loss”.

Corrected.

Lines 43-51: Suggest emphasizing in this paragraph that the dates represent a time-integrated thermal history. This paragraph focuses only on simple exhumation scenarios, which differ from the deep-time studies discussed in this paper, that include multiple burial and erosion events.

Corrected.

Line 57: suggest rewording to “the transition from open-system to closed-system” given the phrasing of the previous sentences.

We have kept this sentence because it is quite general.

Line 59: Appropriate to cite Ketcham et al. (2013) here, along with Guenthner et al., 2013.

Cited.

Line 67. “Both those in favour of a glacial (McDannell et al., 2022) and tectonic (Flowers et al., 2020) origin…”

To be more accurate, suggest modifying to “Both those in favor of a globally synchronous glacial origin (McDannell et al., 2022) and regionally diachronous tectonic origins (Flowers et al., 2020)…” The text has framed the problem here and elsewhere as a debate over a single origin for this feature – but a critical aspect of the debate is whether or not there is indeed a singular origin or if there are multiple origins.

Corrected.

Lines 69-77. This is a great description, and also nicely highlights that annealing at higher temperatures after damage accumulation at low temperatures affects the retentivity and the date. It's this element that could be useful to bring into the geological framing in lines 43-51.

Thank you! We want to keep the general structure and keep the intro to thermochron as concise as possible. With all the shifting of the intro, we have hopefully achieved this and feel like it is acceptable to focus on the problem here.

Line 79. Suggest changing RDAAM to ZRDAAM, as ZRDAAM was defined in previous paragraph. RDAAM refers to a widely used apatite radiation damage accumulation and annealing model, which isn't discussed in this paper.

Corrected.

Lines 81-85. "Parts of this history were reported to within less than 10 degrees between 700 and 250 Ma and then again from 15-7 Ma. It is unclear whether the data really provide such tight constraints on temperatures in the past…"
- It's clear that the data don't provide such tight constraints on the temperatures in the past, because the model of their Figure 4 says that when the Great Unconformity developed that the basement was at 200 +/- 10 °C (so, 7-10 km deep in the crust, depending on assumptions), when we know unequivocally that the rocks were at the surface during deposition of the Cambrian sandstone on the basement to define the Great Unconformity. This model isn't set up to require that the Great Unconformity be part of the model. See Peak et al., 2022. This doesn't seem like a good example to use here.

We have kept this example here to highlight that sometimes thermal histories appear to be too well resovled. Because our focus is on the importance of damage model uncertainties, this is a really nice example to highlight as our damage model would increase the uncertainties.

Lines 87-88. "…known amounts of radiation damage." Suggest being more cautious here. Instead of saying "known amounts of damage", could say "…well-constrained amounts of radiation damage." For the Sri Lankan zircon, which I agree was a good sample to use in the Guenthner et al. study, the assumption of rapid cooling at 440-420 Ma to estimate radiation damage seems reasonable, but this is a long timescale with some uncertainty in the history. And even if one acquires Raman data, there is currently ambiguity in how to translate that data into the type(s) of damage in the crystal that matter for He diffusion.

Lines 94-95. "We show that natural variability in radiation damage annealing parameters causes ZHe ages to be dispersed even for crystals…" Delete "annealing"?

As I understand it, this paper doesn't address variability and uncertainty in radiation damage annealing kinetics?

Corrected.

The existing calibration of the radiation damage and annealing model
Line 123. "…at a specific cooling rate." And assumed grain size.

Corrected.

Line 139. "..between the model parameters." Helpful to specify here that you mean zDo and zEa.

Corrected.

Lines 143-144. "However the accuracy of this model has only been assessed by looking at general trends in model predictions." Do you mean in Guenthner et al.? Or in the variety of studies that have aimed to explain data using this kinetic model? Provide some references here?

We have only cited the Guenthner paper here as these data are the best to constrain the parameters and this is the only time the model has been compared to the data in a quantitative way.

Lines 125-146. Clear explanation of the Guenthner model.

Thank you!

It would be worth explicitly noting somewhere that this contribution does not address uncertainties in the annealing model.

We have highlighted this point at the end of the paragraph. Lines 144.

A new calibration of the zircon radiation damage and annealing model
This section is the crux of the paper and I think that more explanation is needed. I believe that I understand the Guenthner calibration, but I don't understand how the end member values for this new calibration (and thus their associated uncertainties) were obtained here based on the explanation provided. If what is presented here is indeed an improved approach as the authors are clearly arguing, then this is exciting and it should be explained with enough clarity so that those who generate the datasets used for model calibrations can understand and apply this method. If the authors would prefer not to lengthen the main text, then the appendix would be a great place for an extended description.
Lines 194-196. From the Figure 1A and 1B plots, it's not obvious that the revised calibration predicts the observed data better than the original calibration. It would be helpful to add some text here that explains what features of this plot that the reader should focus on to see this.

We have refined this plot by adding colour that reflects total alpha dose. We have also stressed that both models do not fit all the variability in the data. Our aim to assess the uncertainties, not provide a new damage model.

Figure 1. If I understand correctly, the purpose of this figure is to show how well the Arrhenius data for each sample used in the ZRDDAM calibration are fit with the original calibration and the ZRDAAM calibration of this paper. Right now, it's difficult to make this comparison across the two plots. This probably could be done more successfully by including a separate Arrhenius plot for each sample, labelling each Arrhenius plot with the sample name, and on the same plot including the diffusion kinetic data, the prediction from the original ZRDAAM calibration, and the prediction from the ZRDAAM calibration for that sample.

This is not the goal of our analysis. It would also make too many figures and complicate the simple message we wish to highlight. This would be the core of a new paper that attempts to improve the fit the to the data with a new parameterization.

Lines 196-203 and Figure 2. Could you please explain more fully and clearly how these histograms were calculated. Again, this is the crux of the paper, since this calibration and the associated uncertainties are then applied to draw the main conclusions of this contribution.

We have improved the overall description of the model and the parameters.
Propagating model uncertainties
Lines 222-223. "Note, this implementation of the model has been used previously (Tripathy-Lang et al., 2015). It sounds like this model calibration was presented seven years ago in previous work? To what extent was it described there? This should be explained in the introduction.

We have clarified that this refers to the numerical methods.

Lines 224. "...an effective U concentration ([eU = [U] +0.24[Th]; Gastil et al., 1967))." I hadn't encountered this paper before, which interestingly uses eU to refer to "equivalent uranium" (although not effective uranium). That paper doesn't include an eU equation. Cooperdock et al. (2019) should be cited here for the equation, (eU = U + 0.238Th). Flowers et al., 2022 also lays out the equations for eU.

Added references here.

Line 241. "...and the data from McDannell et al. (2022)". Miltich (2005) should be cited for the data. McDannell et al. mined the Minnesota data from the Miltich (2005) undergraduate honors thesis without publishing them.

Added the references here.

Line 242. "…for the sample "Minnesota"… These are actually multiple samples, not a single sample.

Corrected and clarified.

Lines 241-242 and Figure 5. "For this reason, we use only the ZHe data and do not incorporate additional constraints."
- This simulation is supposed to explain the Great Unconformity, but the highest probability time-temperature paths aren't at surface conditions when the Great Unconformity developed. In the final comparative statement that compares the three panels, the caption states "The overall patterns are very similar, but the apparent resolution is different, resulting in different geological conclusions." To me, the geological conclusion here should be that all of these model results are geologically meaningless because the highest probability tT paths violate the Great Unconformity, so all three models should be discarded. Could the authors either include the Great Unconformity in the models or better articulate why the Great Unconformity isn't honored in a model that is supposed to explain the Great Unconformity? This is now a repeated characteristic of many published QTQt models that are supposed to reproduce the Great Unconformity.

We have redone the analysis and the GU is reproduced to some degree now. The discussion about whether to include constraints or not is at the core of thermal history modelling and requires a separate discussion, not a paragraph in a paper on damage models.

Figure 5. Please plot the data being modeled here and show how well the tT paths fit the observed data – for example, by making date-eU plots of observed vs. modeled data. See Gallagher (2016).

The data are well predicted by the model. These plots can be seen in the original publications.

Implications
Lines 264-266. "In turn, it may be challenging to resolve cooling histories sufficiently to attribute the Grat Unconformity to Cryogenic Glaciations (McDannell et al. 2022) or geodynamic process related to the break-up of Gondwana (Flowers et al., 2020)." Yes, agreed that it's not currently possible to resolve cooling histories at this level with the "thermochronologic data alone" even when not considering uncertainties in kinetic models. This is why it is essential to integrate other types of information into models, such as geologic data. Flowers et al. (2020) didn't argue that we could resolve the cooling histories sufficiently without other information, as implied in this sentence. Perhaps you could modify to remove that implication?

Corrected.

Lines 268-281, and final sentence in this paragraph: "For example, McDannell et al. (2022)'s results for Pikes Peak highlight how models that ignore overdispersion appear

to resolve a 700 Ma cooling signature, which is smoothed out when the overdispersion is effectively reduced by adding excess uncertainty on some of the data."
- There are interesting and important points made in this paragraph about how uncertainties can be accounted for in QTQt and the associated influence on the inferred tT paths. The final sentence that makes a vague reference to a figure in McDannell et al. (2022). It would be great if the authors would add a figure in this paper that helps to illustrate the valuable points being made here rather than vaguely referring to a figure in that paper. Alternatively, this paragraph could be eliminated.

We have kept this to highlight the discussion. We tried drafting a new figure but were unable to make an effective figure.

- The Pikes Peak models in McDannell et al. (2022) also violate the Great Unconformity relationship, as noted elsewhere about other published models.
- If this model is discussed, the I feel that it also is important to cite our 2022 comment on this paper. In that comment we show that entirely different tT paths, not captured in the McDannell et al. models, can explain the data.
Lines 283-314. This is a nice illustration of the dispersion expected in real datasets, and provides some of the rationale for the binning into synthetic grains as done by many who simulate (U-Th)/He data.
Code Availability. It would be appropriate to put the codes used for the calculations in this paper in a supplement and available for download so that others can reproduce these results and apply the approach to other kinetic datasets. This is now done so easily that it no longer seems appropriate to require an email to the authors to obtain the code.

QTQt is not available online. We are also attempting to do this same analysis with different diffusion datasets.

I enjoyed reading this contribution and hope that the authors find these comments helpful to further strengthen the manuscript.

Becky Flowers, CU Boulder

References

Cooperdock, E.H.G., Ketcham, R.A., and Stockli, D.F., 2019, Resolving the effects of 2-D versus 3-D grain measurements on (U-Th)/ He age data and reproducibility: Geochronology, v. 1, p. 17–41, https://doi .org /10 .5194 /gchron-1-17-2019.

Flowers, R.M., Ketcham, R.A., Macdonald, F.A., Siddoway, C.S., Havranek, R.E., 2022, Existing thermochronologic data do not constrain Snowball glacial erosion below the Great Unconformities: Proceedings of the National Academy of Sciences, Letter to the Editor, v. 119, No. 38, https://doi.org/10.1073/pnas.2208451119.

Flowers, R.M., Zeitler, P.K., Danišík, M., Reiners, P.W., Gautheron, C., Ketcham, R.A., Metcalf, J.R., Stockli, D.F., Enkelmann, E., and Brown, R.W., 2022, (U-Th)/He chronology: Part 1. Data, uncertainty, and reporting: Geological Society of America Bulletin special volume on the Reporting and Interpretation of Geochronologic data, https://doi.org/10.1130/B36266.1.

Ketcham, R.A., Guenthner, W.R., and Reiners, P.W., 2013, Geometric analysis of radiation damage connectivity in zircon, and its implications for helium diffusion: The American Mineralogist, v. 98, p. 350–360, https://doi .org /10 .2138 /am .2013 .4249.
L. Miltich, 2005, Low temperature cooling history of Archean gneisses and Paleoproterozoic granites of southwestern Minnesota. Undergraduate thesis, Carleton College, Northfield, MN.
Peak, B.A., Flowers, R.M., Macdonald, F.A., and Cottle, J.M., 2022, Forum, Reply to Comment on: Zircon (U- Th)/He thermochronology reveals pre-Great Unconformity paleotopography in the Grand Canyon region: 50 (3): e544, https://doi.org/10.1130/G49965Y.1.

Reviewer: Willy Guenthner

Our responses are to the reviewer's comments are shown in italics.

The manuscript by Fox and co-authors re-examines certain aspects of a commonly used damage-diffusivity model for the zircon (U-Th)/He system (ZRDAAM). The authors highlight that the model, as originally derived and defined by Guenthner et al. (2013), currently lacks information on kinetic uncertainty, which could in turn influence the thermal history modeling-based outcomes that rely upon ZRDAAM. In particular, the authors argue that the lack of kinetic uncertainty could create over-interpretation of and/or overconfidence in thermal history model output as applied to a recent discussion surrounding the origin of the Great Unconformity erosion surface. I found this contribution to be of importance for three reasons: 1) it stresses the need to incorporate uncertainties into the kinetic models used by many authors and thermal history modeling software packages, 2) it shows how large zircon He datasets (~40 grains) are needed to refine and extract meaningful time-temperature information, and 3) it recognizes that ZRDAAM is a work in progress that could be strengthened by further He diffusion kinetic work. In sum, I think this manuscript is a solid contribution to ongoing discussions surrounding thermal history modeling and thermochronometric data interpretation. However, given the authors stated goals of exploring calibration issues with ZRDAAM, I think the article does not go far enough in investigating multiple aspects and nuances of the current ZRDAAM calibration. My recommendation is major revisions, which should focus on expanding the scope of the current work. I have several general comments that I would like to see the authors address before this article is published.

We thank Willy Guenthner for his detailed review and we have addressed as many points as possible. It is important to note that we are not really interested in producing a new damage model – we are simply interested in propagating uncertainties. Points about changing the end member crystal are useful to consider but this requires a completely different model. Furthermore, the approach we develop here can easily be modified to constrain parameters in any damage model. Our goal is to use the diffusion data directly and not rely on an intermediate regression as this ignores correlations in the data.

It should also be noted that we don't think that the endmember crystal should really ever be a real crystal. What we have done is attempted to fit all the diffusion data with the same model. In this way, all the data contribute to determining the diffusion kinetics of the end member crystals. We think this is preferable to using something like Mud Tank because then the other crystals and diffusion experiments would not be utilised.

First, the authors focus heavily on the theoretically pristine endmember zircon, and the fact that this value is an extrapolation that, correctly, could have a large degree of uncertainty to it. It is difficult in the current preprint to determine how important this endmember is in the context of real datasets. Let me expand. When the ZRDAAM was derived, a proposed "zero-damage" grain was used to make the model setup more straight-forward, and to account for potential scenarios where users may want to model thermal histories at very short time scales. But nearly all grains, especially those used for deep-time problems, will very quickly (within ~5 myr for a grain with 500 ppm eU) accumulate damage levels in excess of the Mud Tank sample, for which empirical results do exist. So perhaps the challenge with the current ZRDAAM is less, what should we do for a zero-damage endmember, but rather, should we instead be using a sample with known kinetics as our endmember? Would the Mud Tank kinetics therefore be more appropriate here as a realistic endmember, and if so, does that reduce some of the variation in thermal history model output? Secondarily, there are both molecular dynamics (Reich et al., 2007; Saadoune et al., 2009) and empirical results from zircon-like, zero-damage orthophosphates (Farley, 2007) that attempt to place an approximation on He diffusion in a pristine zircon. How do the author's new extrapolations match up to these results as a comparison? Finally, as Ginster et al. (2019) observed, it is very unlikely that a natural zircon would return to a pristine damage state following damage accumulation, so the likelihood of encountering a natural, zero-damage zircon is remote. I would like to see the authors acknowledge some of these points and address these questions. I think the authors have illuminated a valid and important point: that the ZRDAAM as defined by Guenthner et al. 2013 needs to be recalibrated at the low damage end and uncertainty needs to be accounted for. However, as I've hopefully conveyed, I would place the focus more upon using an endmember for which we can obtain kinetic information, rather than (as was done in Guenthner et al., 2013 and is done here) an idealized, extrapolated endmember that does not likely exist in nature.

We have discussed this point above because we feel like this is the main point of the discussion. We have also addressed this in the text by reiterating that we are not aiming to produce a new model. We just want to propagate uncertainty.

Second, the authors highlight a data set from Minnesota as an example with which to test their newly constrained ZRDAAM. My concern with this modeling setup, as articulated in a line specific comment below, is that an apples-to-apples comparison to the McDannell et al. 2022 study would include many more iterations. In that study, 500,000 burn-in, and 500,000 post were used, whereas here it appears that only 100,000 burn-in and 100,000 post are used. That is, the higher amount of uncertainty as to when Neoproterozoic cooling initiates could be that the model was not run for long enough.

This is one reason why the models would be different. But there are many other reasons. We are not comparing the models to the McDannell model, but rather the results with different model parameters.

Third, the authors call for additional diffusion experiments and new data to be added to the ZRDAAM, which I certainly agree with. I would point out though that some of these data have already been collected by Ginster and are published in her PhD dissertation, available through multiple repositories, including the University of Arizona library. This work is yet to be published in a peer-reviewed manuscript, but the data essentially agree with the Guenthner et al. (2013) diffusion data, which suggests that the variation (and therefore the uncertainty) in measurable diffusion kinetics is lower than for extrapolated endmembers. The point should be emphasized that, from an experimental perspective, the ZRDAAM remains on solid footing, although perhaps the endmember kinetics need to be recast, as my previous comment suggests.

Totally agree with this point. Again, we just want to propagate uncertainty. However, we don't think it is appropriate to include the data in the inversion for the diffusion kinetics. These data are unpublished and we would rather wait until Ginster publishes her thesis. Of course, we can redo the calibration again.

Line specific comments:

Line 73: Should acknowledge here that alternatives using an arguably more direct measure of alpha dose (Raman spectroscopy) have been proposed (Ginster et al. 2019) and incorporated into ZRDAAM (Guenthner, 2021). I would also elaborate the discussion here to comment on the different styles of damage that may influence He diffusion kinetics (i.e. alpha ejection, alpha recoil, fission track). The diffusion kinetics are calibrated to alpha dose (essentially alpha recoil), but the mode of annealing is debated and as yet not fully resolved. The Ginster et al. (2019) data and the model demonstration by Guenthner (2021) of these data are particularly salient given my comments above concerning an idealized zero-damage endmember.

We have attempted to keep the paper simple and focus on a single message. Already, this paper is quite complicated.

Line 85: An additional explanation here is that, as McDannell et al. 2022 showed, these models needed to be run for many more path iterations (at least 500,000 pre and post burn-in). In the Thurston et al. (2022) study, we ran models for only 100,000 pre and post burn-in, which was admittedly likely not enough. To this point, we have re-run some of these models with the greater number of paths (currently unpublished) and indeed the earlier portions of the time-temperature history remain under-constrained by zircon He (as discussed in Thurston et al., 2022). However, the late Miocene cooling remains robust.

We have added that parameter space might not have been sampled sufficiently.

Line 99: The phrase "varies by hundreds of millions of years" should be qualified here. As you suggest later in the manuscript, the variation is dependent on the number of grain inputs you have and the spread in date-eU space of those inputs. Moreover, my understanding (this could be clearer, see next comment below at line 238) is that this is for scenarios that incorporate the 2 sigma from the full kinetic distribution.

This has been removed.

Line 238: The focus in this paragraph on fixed endmembers seems out of place with one theme of this manuscript: kinetic uncertainties should be sampled in the rjMCMC approach. Some of this could be my confusion: am I correct that the models were run with Ea and D0 values that represent the 2 sigma of the kinetic distribution? If I'm not correct, then please more thoroughly explain how or why these specific kinetics were selected. If I am correct, then I understand that the authors are perhaps trying to show the worst-case, 2 sigma extremes from their distribution, but why not incorporate the full probability distribution as shown in figure 2 and sample that? The authors mention further below that computation limitations prohibit this exercise, but much of the discussion and implications seem to be cast in light of the highest possible amount of variation. If the modeling incorporates the full distribution (and samples it) is the situation as dire? I am really curious to see the outcome of a modified MCMC proposal algorithm with a selection statement that samples the kinetic distribution.

It is not clear to us how to do this at present. We have a produced a new version of QTQt that explicitly samples from our inferred model parameters. In this way the model does not attempt to find the best combination of diffusion parameters – the diffusion experiments are the best way to do this. Instead it incorporates this uncertainty into the analysis.

Line 249: For a better apples-to-apples comparison here with the McDannell et al. (2022) study, 500k pre and 500k post burn-in is needed. As we have seen (and learned, see the comment about Thurston up above) these deep-time problems require at a minimum 1,000,000 path iterations.

We tested this point and ran models with more and fewer iterations in the burn in phase. This doesn't seem to be a big problem if the chain is properly sampling parameter space. Perhaps, the models with long burn ins are getting stuck in local minima?

Line 294: Are the authors suggesting here that the binning and averaging approach has limitations? The point I'm most struck by, as the authors allude to, is that binning and averaging can eliminate the sensitivity of the whole data set by removing portions of date-eU space.

We agree that this is problematic and would suggest not binning and averaging data. Instead, the distribution of age as a function of eU could be mapped out more completely.

Willy Guenthner

UIUC

Citation: https://doi.org/10.5194/gchron-2022-23-RC5

Reviewer: Kip Hodges

Our responses are to the reviewer's comments are shown in italics.

Review of Fox et al., "Origin of Great Unconformity Obscured by Thermochronometric Uncertainty"

We thank Kip for his helpful and insightful review.

In this manuscript, Fox and co-authors present a detailed and useful discussion regarding the effects of uncertainties related to the Zircon Radiation Damage and Annealing Model (ZRDAAM) on thermal history models based on zircon (U-Th)/He data. They go on to suggest that these effects complicate thermochronologic models that have been used to constrain plausible causes for the Great Unconformity at the end of the Precambrian.

After a brief review of the reasons why some sort of ZRDAAM-like model is necessary to reasonably interpret zircon (U-Th)/He data, the authors correctly note that the original ZRDAAM formulation of Guenthner and co-workers (Guenthner et al., 2013), based on helium diffusion experiments on variably damaged zircons, has shortcomings that make it difficult to realistically propagate uncertainties in two experimentally derived helium diffusion parameters – the pre-exponential constant Do and activation energy Ea – into ZRDAAM models and onward into thermal history models. In order to do this more quantitatively, they first present a revised ZRDAAM formulation in Section 3 of the manuscript. They show that diffusion parameters derived from experimental data using the new formulation mostly agrees with the diffusion parameters adopted by Guenthner et al. for a highly damaged ("amorphous") crystal, but differ markedly for an "undamaged"crystal (Figures 1 and 2). They attribute this to difficulties in estimating Do and Ea independently using the approach of Guenthner et al. Most importantly, however the authors show that the uncertainties on derived parameters are quite large, much larger than the thermochronology community generally assumes.

Yes, these uncertainties are quite large. However, the parameters are also strongly correlated. This means that the impact on the estimates is not quite a large as might be expected, if the parameter correlation structure is accounted for.

Using the revised ZRDAAM formulation, Fox et al. then proceed to explore how such large uncertainties in radiation damage and annealing model parameters might affect thermal history models calculated using the QTQt software package of Gallagher et al.

(Gallagher, 2012). Unfortunately, direct propagation of such uncertainties into inverse modeling of thermal histories using QTQt is computationally impractical (lines 302-305), but the authors present the results of efforts using simplified models showing that uncertainties of up to several hundreds of million years in the estimated timing of a cooling event should not be unexpected (!). They go on to point out, specifically, that these results have dramatic implications for studies such as those being conducted by several groups on the Great Unconformity.

Yes, we argued that this has major implications for when the thermal history is resolved. The impact on when specific events are resolved along a thermal history is a bit more complicated. This is highlighted in the comment by Kalin McDannell. We have dealt with this point more clearly in our revised manuscript.

I enjoyed reading this well-written contribution and feel it is an important step toward developing a robust appreciation of how confident we should be in the results of thermal history modeling studies based on datasets for which either zircon or apatite (U-Th)/He radiation damage modeling is required. I found no fault in the mathematics presented here and I trust the authors to have done the modeling carefully. I hope, however, that these findings are not used by non-specialists to conclude that thermal history modeling using thermochronologic data is practically useless because the uncertainties are so large. Fox et al. have tried a bit to guard against that, but they might be more explicit about that in the Implications section. The principal lessons I think we should learn are: 1) we should not overinterpret every wiggle in a modeled time-temperature path as having geologic significance; 2) we should not overestimate the precision with which we know the timing of specific cooling (or reheating) events given the imprecision of derived diffusion and annealing parameters; and 3) we should all continue to work hard to improve our understanding of the impact of radiation damage on diffusion parameters through a combination of experimental and empirical studies. I was happy to see the nod toward also using natural laboratories to address these issues moving forward (lines 319-322). In fact, led by Alyssa Anderson, our research group at Arizona State performed such a study (Anderson et al., 2017), concluding that the Guenthner et al. formulation failed to predict ZHe closure behavior consistent with QTQt models of an independent multichronometer dataset for the McClure Mountain syenite of Colorado. However, Anderson and co-workers preferred to intereprет this inconsistency as a consequence of variable and complex radiation damage zoning in individual McClure Mountain zircons rather than uncertainties in radiation damage. Certainly, however, uncertainties in the original ZRDAAM model might have played a role, as the Fox et al. work indicates. (It should be noted that, while the accuracy of the cooling history we derived has been disputed based on geological interpretations (Weisberg et al., 2018), the alternative proposed cooling history still remains inconsistent with the ZrnHe closure behavior predicted by the original ZRDAAM model (Anderson et al., 2018).) Our study and the current manuscript together remind us that the effects of radiation damage on ZHe closure behavior of ZHe is extremely complex and dependent on both the parameters we choose for modeling and the idiosyncracies of specific zircons. In such a climate of uncertainty, interpretive caution is advised.

We have included these citation in our revised manuscript and thank Kip for highlighting these to us. It is also nice to see the three lessons we should learn so clearly articulated. This is exactly what we hope people take from the analysis.

– Kip Hodges

References

Anderson, A. J., Hodges, K. V., and van Soest, M. C., 2017, Empirical constraints on the effects of radiation damage on helium diffusion in zircon: Geochimica et Cosmochimica Acta, v. 218, p. 308-322.

Anderson, A. J., Hodges, K. V., and van Soest, M. C., 2018, Comment on 'Distinguishing slow cooling versus multiphase cooling and heating in zircon and apatite (U-Th)/He datasets: The case of the McClure Mountain syenite standard' by Weisberg, Metcalf, and Flowers: Chemical Geology, v. 498, p. 150-152.

Gallagher, K., 2012, Transdimensional inverse thermal history modeling for quantitative thermochronology: Journal of Geophysical Research, v. 117, p. 2156-2202.

Guenthner, W. R., Reiners, P. W., Ketcham, R. A., Nasdala, L., and Giester, G., 2013, Helium diffusion in natural zircon: Radiation damage, anisotropy, and the interpretation of zircon (U-Th)/He thermochronology: American Journal of Science, v. 313, no. 3, p. 145- 198.

Weisberg, W. R., Metcalf, J. R., and Flowers, R. M., 2018, Distinguishing slow cooling versus multiphase cooling and heating in zircon and apatite (U-Th)/He datasets: the case of the McClure Mountain syenite standard: Chemical Geology, v. 485, p. 90-99.

Reviewer: Brenhin Keller

These responses follow the same structure as the original review. We do not include it here because the review was so thorough.

We thank Brenhin Keller for his helpful comments on the manuscript. We would also like to apologise for misrepresenting previous work. We fully appreciate that the authors' work went far beyond simply trying to infer a glacial origin of the Great Unconformity using a single sample analysed with a single thermochronometric system. It is clear that everyone appreciates that multiple samples analysed with multiple thermochronometric systems is the best option. We have clarified this here and have modified the text to highlight this. Furthermore, we appreciate that a glacial origin would not mean synchronous erosion at all locations because glacial erosion is complex. We have modified the text accordingly. We decided to use the GU discussion as a focus for our analysis because it is an exciting problem in Deep Time thermochronometry where uncertainites are likely to have an important impact on the results.

We have modified the title to ensure that is it more general and reflects the fact that damage models are important for a host of different applications. This title better reflects the content of the manuscript but maintains a focus on the GU because this is such an exciting field.

With respect to the issues highlighted with the QTQt modelling, we attempted the model with different burn in durations and found no changes to the solution. It is not clear why the burn in length has such a clear control on the models presented by Kalin McDannell in his comment. All QTQt models were carried out by Kerry Gallagher and the acceptance criteria were carefully monitored. The linear cooling models also look a lot like the maximum posterior solution, even in the updated versions of the plots. These models are probably very sensitive to the size of the prior, the acceptance rates and the proposal distributions. Identifying the "correct" or most suitable thermal, and its uncertainty, is far beyond the scope of our study.

We were able to reproduce the overall inferred thermal history by decreasing the uncertainty in the ages to 25% of the age. This provides us with a relatively precise thermal history and we can see how this changes as model uncertainty is increased. In this way we are expanding the errors but by propagating data uncertainties. Inferred histories from different datasets will be impacted differently by these expanded uncertainties. This exploration should be the focus of a different analysis and we make no recommendations here or suggest that a thermal history is better than any other. We are simply interesting in how uncertainty in the diffusion parameters influences uncertainty in the thermal models. We have developed a method to do this by modelling the diffusion data directly and propagating this uncertainty directly to the QTQt models.

We have introduced a way to quantify uncertainty in radiation damage models and propagate this through to thermal history uncertainty. We have not attempted to estimate uncertainty in radiation damage annealing, which would also influence the inferred thermal histories. For this reason, we would like to reinforce the point raised by Brenhin Keller and other authors, that incorporating additional datasets is important. Incorporating additional thermochronometric data is always a good thing. Incorporating geological constraints is also useful, but it is also interesting to assess what the data can resolve without independent constraints. Ultimately, by modelling the data without constraints, we can investigate what additional data is needed to get back to the correct constraints.

Reviewer: Kalin McDannell

We appreciate the detailed comment provided by Kalin McDannell on our paper. We have reframed some of the introduction and key points in our manuscript to highlight that this is indeed a general problem and certainly not unique to deep time thermochronometry. We would like to take this opportunity to clarify some points, but many of these points should really be discussed in a separate paper that goes into more detail. We have broken this response down into the same sections as the original review.

1. Kinetic model uncertainty

We appreciate that the uncertainties that we have highlighted are not unique to the Great Unconformity debate. We have added text to highlight that this is important for other applications. We have also restructured the introduction to begin with a section on general thermochronometric methods. However, much of the work on deep time thermochronology relies on interpreting data that have been influenced by radiation damage. This is why we focused on this interesting point.

2. Parameter correlations

We think we have done a better job at accounting for these parameter correlations in our revised manuscript. Now we have a modified version of QTQt that explicitly accounts for uncertainties in the damage model. We have also approximated the distribution as 4D guassian that allows us to explicitly define a covariance and correlation matrix. These clearly highlight the model correlations.

3 .The effects of additional constraints on inversions results?

We have removed these inversions because they have caused so much controversy and detract from our very simple message. Instead we show a new QTQt model where we have reduced the uncertainties on the ages. We were unable to replicate the thermal histories inferred from in the comment using the reported data uncertainties. We modified the length of the burn in but this had no effect on the inferred thermal history. This work was carried out by Kerry Gallagher, who is an expert in thermal history modelling.

4. Posterior probabilities

See previous point.

4.1 MCMC burn-in and ZHe uncertainties

K. Gallagher ran the QTQt models with the quoted errors from the original publications. Changing the burn in length made no appreciable difference. Instead, we think the issue may be related to the data uncertainty. We have attempted to address this here and we are now able to recreate a GU-type thermal history.

4.2 Timing of cooling

It is not clear what the reviewer is referring to here. We had used the point in time at which the green colour first appears as this represents when the path is well resolved. This is when the thermal history is resolved, we are saying nothing about when a period of cooling is resolved.

4.3 Continuous spread of ages as a function of eU

Yes, we would argue against binning data in general. The binning of the data potentially removes valuable information that might guide future research. However, binning may be required in certain cases.

5. Bayesian MCMC inversion tests

These models are very interesting and highlight the importance of testing model parameter values. We thank Kalin for highlighting this and we have gone to great lengths to fully explore this in the new QTQt section of our paper.