



EGUsphere, author comment AC1  
<https://doi.org/10.5194/egusphere-2022-299-AC1>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC1

Peter E. Martin et al.

---

Author comment on "Calculation of uncertainty in the (U-Th)/He system" by Peter E. Martin et al., EGU sphere, <https://doi.org/10.5194/egusphere-2022-299-AC1>, 2022

---

We thank Prof. Vermeesch for taking the time to write a prompt, thoughtful, and helpful review. While it will become clear that we largely disagree with the conclusions he draws about the content of the paper, he has helped us identify several areas where the manuscript can be improved and clarified. Below, we provide a high-level response that addresses many of the reviewer's critiques. We then respond to the individual points raised, copying the text of the original review in bold.

Much of this review centers on shortening the text and streamlining the equations provided to yield the most compact possible description of HeCalc and uncertainty in the (U-Th)/He system. In short, we wish to write this paper for the benefit of all geoscientists, not just those with a strong grounding in math and statistics. While the reviewer likely believes that our approach to this goal largely underestimates readers, we disagree. Our experience in the CU TRaIL (Thermochronology Research and Instrumentation Lab) with hundreds of researchers, students, and other clients has demonstrated that the levels of comfort with statistics and math within the (U-Th)/He user community are highly variable. Our objective is to make the manuscript as comprehensible to as wide a readership as possible in an attempt to subvert the well-known "Curse of Expertise" wherein knowledge experts underestimate the difficulty of a subject for novice and intermediate practitioners (e.g., Hinds, 1999; <https://doi.org/10.1037/1076-898X.5.2.205>).

The fundamental purpose of this paper is not to simply publish a method for (U-Th)/He uncertainty propagation, but rather to fully explicate this process for the full range of practitioners of this method, from lab managers and PIs to external users of (U-Th)/He data and early graduate students. Shortening the manuscript in the manner suggested by the reviewer would remove a significant amount of text dedicated to not just writing out the formulae for (U-Th)/He uncertainty propagation, but meant also to assist readers in developing an intuition for this system. We now recognize that this primary goal of accessibility is not well communicated in the text. We will correct this and thank Prof. Vermeesch for highlighting this shortcoming.

**This paper discusses the error propagation of (U-Th)/He data. Its introduction claims that "the formal analytical uncertainty in (U-Th)/He dates has never been thoroughly assessed". I have three comments about this statement:**

The full statement in the paper is "methods for propagating uncertainty components into single-grain (U-Th)/He dates have never been described in the literature, and the formal

analytical uncertainty in (U-Th)/He dates has never been thoroughly assessed.” We largely stand by this statement, emphasizing that we are not claiming that it has never been done, but rather that it has not been fully explained in the literature, as we intend to do in this manuscript. Again, we will revise the manuscript to more clearly communicate the latter goal.

- **I am sure that the error propagation of the (U-Th)/He method has been worked out before, and probably several times. In fact, I have done so myself, and even implemented it in a publicly accessible computer program: <https://ucl.ac.uk/~ucfbpve/heliocalc/>.**

It is true that we failed to mention that a tool for performing this uncertainty propagation is presently available. However, the methods by which this program functions are not (to our knowledge) available outside of inspection of the source code. As this source code is largely uncommented and is written in javascript (in our experience, an uncommon language for geoscience computing compared with Fortran, Matlab, Python, R, or Julia), we believe it would be difficult for a non-expert to understand how or why the program produces the uncertainties that it does. Again, it is the inaccessibility of the knowledge behind uncertainty propagation that we also wish to address, so that it is no longer a “black box” to users.

- **One probable reason why nobody published the error propagation formulas for the (U-Th)/He method is the overdispersion that characterises most (U-Th)/He datasets: the scatter of several aliquots from the same samples usually exceeds the precision of the data, by a lot. So, in a sense, the analytical uncertainties are irrelevant. The interplay between analytical uncertainty and overdispersion is discussed by Vermeesch (2010, doi:10.1016/j.chemgeo.2010.01.002), who also covers some aspects of the error propagation problem.**

We strongly disagree with this line of reasoning. This argument may have been appropriate to make 15 years ago, but the field of (U-Th)/He dating has advanced radically since then. Overdispersion is a central motivation for this contribution. How can we understand overdispersion if we don't first carefully propagate the uncertainty in (U-Th)/He dates that can be well-quantified? By first constraining and then subtracting the scatter in data attributable purely to analytical error, one may begin to approach the problem of the underlying physical causes of overdispersion with greater confidence. This phenomenon is one of the core motivations for the work, as stated in the introduction, background, and conclusion sections. We will edit to make this point even more clearly in the text.

- **A second reason why error propagation hasn't been discussed much is that it is next to impossible to quantify the analytical uncertainty of the alpha ejection correction, which is one of the main sources of uncertainty in (U-Th)/He dating. Geochronologists slap a nominal uncertainty on this correction, which largely defeats the purpose of rigorous error propagation for the other variables. Unfortunately, the paper under consideration does not address this issue.**

This claim that it is “next to impossible to quantify the analytical uncertainty of the alpha ejection correction” is simply untrue. Recent (Cooperdock et al., 2019; <https://doi.org/10.5194/gchron-1-17-2019>) and ongoing (Zeigler et al., 2021; DOI: 10.1002/essoar.10507962.1) work is quantifying the geometric uncertainties associated with alpha-ejection corrections, as is summarized briefly in our paper. It also is possible to quantify the uncertainty introduced by zonation (e.g., Hourigan et al., 2005 DOI: 10.1016/j.gca.2005.01.024; Johnstone et al., 2013 DOI: 10.1016/j.gca.2013.01.004).

These considerations are fully discussed in the recent review paper by Flowers et al. (2022; <https://doi.org/10.1130/B36266.1>), in which numerous labs agree for the need to more rigorously quantify uncertainties in alpha-ejection corrections and propagate them into the reported uncertainty in (U-Th)/He dates, rather than “slapping a nominal uncertainty on the correction”. In this manuscript we both provide a method by which to do Ft uncertainty propagation and also fully explain it.

While it is true that in this paper we apply some nominal uncertainties to the  $F_T$  values in Section 5, these uncertainties are meant to illustrate exactly the point the reviewer is making here: that Ft uncertainty is likely a major contributor to intra-sample dispersion, and that quantification of these uncertainties should be a high priority for the (U-Th)/He community.

**Despite these three caveats, I do not object to publishing the error propagation formulas in GChron. However, before this can happen the manuscript needs serious revision. The paper is far too long and can be shortened by at least 50%. I will make some specific suggestions for this later in this review.**

**The paper uses both standard error propagation and Monte Carlo (MC) simulation. I have two comments about this:**

- **According to the authors, the main advantage of the MC method is its ability to handle skewed error distributions. However, it would be easy to adjust the conventional error propagation to handle the observed skewness. This can be achieved by reformulating the error propagation formula in terms of the log of the variables (e.g., Section 5 of <https://doi.org/10.5194/gchron-2-119-2020>). Thus, the log of the age could be calculated as a function of the U, Th and He concentrations. An even better solution would be to use log-ratios. See Vermeesch (2010) for details. I am not sure how easy it would be to reformulate the paper and HeCalc code in terms of log(ratio) variables. If the authors find it too difficult, then I guess that the MC approach would be fine as an alternative.**

We will look into this possibility; an explicit calculation of skewed uncertainties would certainly be a useful addition. On the other hand, we do not believe that the MC approach has any major downsides given HeCalc’s computational efficiency. An added benefit of MC is in its mathematical simplicity, furthering the overall purpose of this paper, which is to provide an easily comprehensible description of the means of deriving uncertainty in (U-Th)/He dates.

- **The actual main advantage of MC error propagation is not mentioned, namely its ability to handle non-Gaussian error distributions. This is particularly pertinent with regards to the alpha-ejection correction (i.e. the uncertainty of the alpha-retention factor  $F_t$ ). Meesters and Dunai (2002, [https://doi.org/10.1016/S0009-2541\(01\)00423-5](https://doi.org/10.1016/S0009-2541(01)00423-5)) and Hourigan (2005, <https://doi.org/10.1016/j.gca.2005.01.024>) have shown that compositional zoning can strongly affect the fraction of ejected alpha particles. Matters are further complicated in the presence of broken grains, when the alpha ejection correction may result in overcorrection (Brown et al., 2013, <https://doi.org/10.1016/j.gca.2013.05.041>). Things are even more difficult for slowly cooled samples, in which alpha-ejection occurs synchronous with diffusive loss of helium. The dispersion caused by all these complexities is difficult to ascertain, but is likely non-Gaussian. The MC approach could be used to explore these effects. I’m not saying that the authors should do this in their paper (because I don’t want to make it even longer), but they should at least mention the possibility. Perhaps HeCalc could offer an interface to**

## **explore these effects?**

This is a good point, and one we had not fully considered. It would certainly be possible to offer a means of inputting additional non-gaussian uncertainty components of arbitrary shape into HeCalc. We will evaluate the feasibility of including it in HeCalc for the final version of this paper, and, if it is not feasible, try to include it in future HeCalc versions.

## **Detailed comments:**

**Equations 4-10 are unnecessary. They are simply repeating Meesters and Dunai (2005, <https://doi.org/10.1029/2004GC000834>), and Raphson and Newton (~1711). Incidentally, I do not really see the point of using the Meesters and Dunai (2005) solution as a starting point for a Raphson-Newton algorithm anyway. Their direct solution is accurate to better than 0.1% for ages up to 500Ma, which covers all terrestrial applications of the (U-Th)/He method.**

Equation 4 is an alteration to the Meesters and Dunai (2005) method that is not included in the original paper, so we think it should probably be kept. The other equations are indeed a reproduction of other works. We included these for complete traceability of the math in the paper, but will defer to the editor on decisions about whether to replicate equations that are available elsewhere.

Many (U-Th)/He dates exceed 500 Ma, including dates produced regularly by CU TRaIL. "Deep time thermochronology" is becoming increasingly popular. Moreover, is there any reason to knowingly accept errors in date computation? The Newton-Raphson method is efficient enough that for ages <500 Ma virtually no loss of computational efficiency is observed. Computational efficiency only suffers for older dates where the Meesters and Dunai method provides poorer estimates, but in these cases the Meesters and Dunai-generated age has greater error anyhow. We see no reason why labs would not report dates using iteration; this approach is favored by labs in the recent GSAB paper on reporting (U-Th)/He data (Flowers et al., 2022; <https://doi.org/10.1130/B36266.1>).

**Equations 11 and 12 could be written more succinctly in matrix form.**

**Equations 14-18 all share the same denominator (which equals  $df/dt$ ), which could be stored in a variable. All these equations could be put together into a single Jacobian matrix, or moved into the appendix.**

Here again, we favor clarity over brevity and feel that the way these equations are written are more widely accessible than their admittedly more compact matrix counterparts. The main reason for including this series of equations is to provide the means by which a lab that does not wish to incorporate HeCalc into their workflow may perform rigorous uncertainty propagation in a spreadsheet program. While it is true that Excel can handle matrix math, this process requires multiple rows which will generally not be compatible with most labs' workflows.

**Section 3.3.2 describes a method to choose the optimal number of MC iterations to derive a desired level of precision on the mean value. It just presents the well known "the square root of n" phenomenon, which I think is too trivial a result to occupy so much space (Figure 2 is certainly not necessary). It is also important to note that the square root of n rule only applies to the standard error of the mean. The standard error of the standard deviation ( $s$ ) is given by  $s/\sqrt{2n-2}$ . I am mentioning this here because the uncertainty of the standard deviation is more relevant than that of the mean, which is never used in the remainder of the paper.**

This is an entirely valid criticism. We will plan to convert the algorithm to select the number of Monte Carlo iterations to a calculation of standard error of the standard deviation and trim this section significantly.

**I installed HeCalc on my computer and am happy to confirm that it works. I have not extensively tested it though. I think that the presentation of HeCalc should take greater prominence in the paper. Of course, this will automatically happen if some of the remaining bulk is removed.**

It is good to hear that an outside user has successfully downloaded and run HeCalc. The usability of new software is always challenging to assess as we are sure the reviewer is aware. As far as “taking further prominence”—is there a component of the description of HeCalc that is missing?

**HeCalc requires that the user provide the uncertainties of the alpha-retention factors 238Ft, 235Ft, 232Ft and 147Ft. However, the paper does not explain how these uncertainties should be obtained. A nominal 5% uncertainty is used in later examples, without proper justification.**

As mentioned above, the means to derive Ft uncertainties is an active area of research and would be beyond the scope of this paper; the 5% Ft uncertainty (and 2%, which was also included) was meant to be illustrative. These values are, however, not arbitrary but are based on recent work (cited in the manuscript) and are a likely reasonable magnitude for the purposes of this study. On the other hand, it is entirely possible to input zero uncertainty for each and obtain the same result as if the uncertainties were ignored by the program.

**HeCalc also requires that the user specify the error correlations between the different parameters. However, it does not discuss how to estimate those correlations. Does the CU TRaIL database specify them?**

Unlike Ft uncertainties, error correlations are not required inputs; the program defaults to assuming uncorrelated uncertainty if these columns are not present. CU TRaIL does not currently include correlations in uncertainty for any regularly calculated uncertainty (i.e., He, U, Th, or Sm)—our methods should in theory result in fully independent error. As discussed in the background section, uncertainties in Ft are likely highly correlated, though we will leave the quantification of this to future researchers. We apply both fully correlated and uncorrelated uncertainties in Section 5 to circumvent this currently underconstrained problem.

**Minor comment: the paper (and HeCalc) use the awkward convention to report MC uncertainties as “68% confidence intervals”. I understand where this comes from: a 1-sigma interval around the mean of a normal distribution covers 68% of that distribution. However, uncertainties are usually reported either as standard errors or as 95% confidence intervals. If the authors want to compare their analytical results with the MC simulations, then a 95% confidence interval would be more elegant.**

This is another good point. Ideally, the program would provide both 1- and 2-sigma uncertainties and their equivalents for non-gaussian distributions calculated by MC to allow apples-to-apples comparison of whatever uncertainty statistic an end user prefers. We will implement both a 68% and 95% confidence interval output for the code.

**Section 5 can be nearly completely removed. The most interesting part of this section is the finding that parent concentrations are a greater contributor to the uncertainty budget than the helium concentrations. This finding could be**

**reported much more succinctly.**

This is where our most important disagreement with this review becomes apparent. Section 5 is almost entirely devoted to a thorough exploration of the mathematical systems described in earlier sections in largely non-mathematical terms, with the intent of elucidating the system for all users. Although this section likely feels boring and repetitive to those with significant experience with statistics, for many, this portion of the paper may be crucial to developing a real and deep understanding of the uncertainty components in a (U-Th)/He date and how they combine to provide a date uncertainty. We feel that the subsection 5.4 in particular is essential. For users with more hands-on experience with real data than statistical/mathematical theory, seeing how the concepts described early in the paper apply to real data may be critical.

**According to lines 421-423: "when combining uncertainties with equal magnitude, the resulting uncertainty will be only ~1.4 times larger than the input, rather than twice as large as might be expected." Here the authors underestimate the reader. I am certain that the vast majority of geochronologists are familiar with the quadratic addition of uncertainties. Consequently this sentence, as well as the preceding paragraph and Figure 4, can be safely removed.**

Though we agree all researchers would be capable of calculating uncertainty by adding components in quadrature and observing the non-linear effects of this approach, in several conversations we have found that users of (U-Th)/He data at a variety of levels did not intuitively and instinctively expect this relationship. Again, our primary goal is increased accessibility to a broad group of (U-Th)/He data producers, users, and students, not the handful of statistical experts who can easily skip over a couple dozen words.

**The paper attributes the reduction of analytical uncertainty with increasing date to the "roll over" of the exponential decay function. This may be correct but is largely irrelevant to real world applications. The observed reduction only expresses itself at >1 Ga, while the vast majority of published (U-Th)/He dates are <200 Ma. At young ages, the helium age equation is linear to a good approximation (<https://doi.org/10.1016/j.chemgeo.2008.01.027>). Note, however, that the fixed uncertainty of the helium measurements shown in Figure 3 is not realistic: older samples will tend to contain more helium, which can be measured more precisely. This will also cause a reduction of analytical uncertainty, even for Cenozoic samples.**

As discussed earlier, ancient (U-Th)/He dates are becoming common as deep time thermochronology gains prominence; dates where this uncertainty reduction becomes important do exist, and are increasingly routine.

Figure 3 demonstrates the relationship between input and date uncertainty, but is not meant to reflect a series of real data (i.e., we are not suggesting that uncertainty in helium measurement will be the same across all dates). Because the shape of this reduction depends only on Th/U ratio, the main use for this figure would be to read off a % reduction in uncertainty at a given date for a given Th/U ratio.

Additionally, the decrease in He uncertainties with sample age is neither continuous nor universal. Many analysts will specifically target grains with a range of U and Th contents to explore the effects of radiation damage, meaning grains with nominally the same date can have very different He contents. Even with increasing He contents, uncertainties do not continuously decrease, but instead are a product of uncertainty resulting from blank and standard measurements, creating a minimum uncertainty bound. It is not the case that a 1000 Ma grain, for example, necessarily has a lower He uncertainty than a 100 Ma grain.

**In section 5.2, the authors introduce a new definition for skewness. This is a very bad idea. There already exists enough confusion in the geological community about basic statistical concepts. It would be unwise to add to the confusion by redefining widely accepted terms such as skewness. At this point I would like to reiterate the fact that the approximately lognormal uncertainty distribution of the dates could easily be captured analytically by recasting the equations as a function of the log of the age. Simply referring to the percent uncertainty of the age would capture the uncertainty and the skewness with a single number.**

This seems like another instance where style preference comes to the fore. Skewness is a widely accepted term within the statistics community, and it is likely that many geochemists are obliquely familiar. We believe it unlikely, however, that most geochemists intuitively grasp the inherent meaning behind the numerical calculation of skew (e.g., what does a skew of +0.8 mean for a given distribution?). The redefinition we include is meant only to assist in reader and user comprehension of the skew of the distributions. If the reviewer thinks it would be appropriate to provide the formal skewness in addition, this would be straightforward to implement, but may unnecessarily clutter the output, especially with the inclusion of 95% confidence intervals.

**Section 5.4 applies the algorithms to a database of ~3,600 (U-Th)/He dates. It is a shame that this database is not released along with the paper. It must be a treasure trove of useful information! Unfortunately, I don't think that Section 5.4 is particularly interesting. It definitely doesn't deserve seven manuscript pages, four pages and three figures (not counting sub-panels). However, Figure 11 does illustrate my comment at the start of this review effectively: the nominal uncertainty of the alpha-ejection correction dwarfs the other uncertainties, thereby defeating the purpose of the careful error propagation.**

Much of the data in this compilation comes from samples CU TRaIL was contracted to run. We do not "own" it and it therefore is not ours to release outside of these anonymized and derived figures. Much of the data produced by CU TRaIL for internal research projects is published or is in the process of being published, and is easily discoverable in the literature. Our lab has played a lead role in developing community agreement on (U-Th)/He data reporting protocols and in making the information needed to understand and interpret (U-Th)/He data accessible to the community (Flowers et al., 2022a,b). However, we cannot release unpublished data without their scientific context and with no means for proper attribution, especially when those data are not ours to release in the first place.

As stated earlier, we feel that this section is in fact highly important as a means of making the results of this study more broadly comprehensible. This final section provides examples and (hopefully) builds intuition for non-experts as to how this system will actually function in practice. Presented with the bare facts in Sections 2-4, we believe many (U-Th)/He practitioners would have the immediate thought "Ok, so how does this apply to my data?" Our goal with Section 5 is to provide an answer to that question. We will, however, revisit with an eye to tightening.

Figure 11 is important precisely to make the point that the current efforts to more rigorously quantify Ft uncertainties are well-worth the time. And at the point that such uncertainties are well-quantified, then they should be propagated into the uncertainties reported on (U-Th)/He dates.

**Lines 615-616: "a challenge to interpreting data with asymmetrical uncertainties is that no widely used inverse thermal history modeling software for (U-Th)/He data permits the input of asymmetrical uncertainty" I'm not sure how HeFTy handles the analytical uncertainty of (U-Th)/He data, but if I seem to recall that**

**QTQt essentially inflates the uncertainties until they account for the overdispersion of the data. This means that the uncertainties are, effectively, ignored. HeFTy probably does something similar, because otherwise its formalised hypothesis tests would fail. Ideally, thermal history inversions should aim to predict the uncorrected (U-Th)/He dates, ignoring the alpha ejection correction. As mentioned before, this is because alpha ejection occurs concurrently with thermal diffusion. So it is not a constant but a variable that depends on the thermal history (Meesters and Dunai, 2002).**

For HeFTy, we are certain that the program takes only the uncertainty provided by the user, which is assumed to be gaussian in the formulation of HeFTy's goodness-of-fit parameter. This issue is explored in Vermeesch and Tian (2014; 10.1016/j.earscirev.2014.09.010) and the subsequent comments and replies, where the authors note that improved uncertainty results in fewer "good" and "acceptable" paths, which is really just an artifact of the random Monte Carlo path generation method combined with tighter goodness-of-fit requirements of improved uncertainty. We are less familiar with QTQt, but it must apply some version of uncertainty to the data. Unless this uncertainty is permitted to be asymmetrical, the overall point stands: date probability distributions have the potential to be non-gaussian, and that potential should be included in thermal modeling.

**Equations a1-a10 all have the same denominator. Storing this denominator in a variable would avoid a lot of duplicate text. You could then even put all these equations into a single concise Jacobian.**

We have the same response to this as with earlier suggestions that would streamline the equations in this paper. Adding variables would just increase the complexity of interpreting the equations. In contrast, the repeated text does not harm the paper in any way (it doesn't even add lines to the manuscript). It is true that these could be recast in matrix math, but for input to an excel workbook, readers would need to revert that change. On the other hand, any practitioner who wishes to convert these equations to matrix form before using them is likely to be more than capable of doing so.

**I apologise if this review comes across as overly critical. I think that this paper (and the HeCalc program) could serve a useful purpose. My opinion is that it would be greatly improved by trimming it down to the important parts. Perhaps the paper could be recast as one of GChron's popular "Technical notes"? This would provide a nice way to present HeCalc to the world, whilst reviewing the error propagation problem.**