

## Authors' response to the reviews

We thank both reviewers for their detailed comments and suggestions. Please find below our response to each comment. Changes made to the original manuscript are indicated in green.

### Authors' response to Review 1

- Reviewer: Olga Yakubovich
- Citation: <https://doi.org/10.5194/egusphere-2025-3879-RC1>

Note: The reviewer comments are presented in **black bold print** in their original, unabridged wording. The author's response is displayed in blue.

**The submitted manuscript is well written and presents a new methodological approach and algorithm for in situ (U–Th–Sm)/He dating of apatite aimed at reconstructing thermal histories of individual grains. Using apatite samples from southern Germany, the authors demonstrate that the proposed methodology enables the derivation of “He diffusion” profiles. It is shown that profiles corrected for parent nuclide distribution are flat in rapidly cooled grain, whereas in samples with heterogeneous parent nuclide distribution and complex thermochronological history, they exhibit non-flat geometries. It is shown how the obtained analytical data can be mathematically processed to derive information on the thermal history of the grains.**

**Thus, I consider this study to represent a valuable advancement in the development of in situ (U–Th–Sm)/He thermochronology.**

We thank the reviewer for their time and careful review of our manuscript. The reviewer raises the following key concerns in their general comments below: Firstly, they state that for one of the apatites we analysed, we did not succeed in deriving a thermochronological path matching the observations in that grain. Secondly, they suggest that analysing well-characterised standard material would have been preferable to our choice of sample material. Lastly, the reviewer notes that some of our methodological approaches require a more detailed description. We address all comments below in detail and are optimistic that our answers and proposed revisions to the manuscript will alleviate the reviewer's concerns.

We have numbered the following reviewer comments (RC) and author responses (AR).

- **RC 1:** My major concern is that the authors did not succeed in deriving a thermochronological path that adequately fits their observations for the BaF apatite. The obtained 4 profiles are reproducible and, as the authors noted, exhibit a counterintuitive geometry: the central part of the grain is significantly younger than the rims.

- **AR 1:** We thank the reviewer for this comment. As they bring up several points at once, we will address each issue separately.
  - On the lack of success for deriving an adequate thermochronological path:  
Regarding the challenges in deriving an adequate thermochronological path for Apatite-BaF, we acknowledge that we were unable to invert all  $^4\text{He}$  profiles of Apatite-BaF for cooling histories. However, we want to emphasise that for profile Ap-BaF-P1, we did obtain fitting cooling histories (see Fig. 6 in our manuscript). We think a major issue with the non-invertible  $^4\text{He}$  profiles is their asymmetry with respect to the c-axis, which conflicts with our inverse modelling approach that assumes approximately c-axis symmetric  $^4\text{He}$  profiles. We will revise Section 4.7 in the discussion to highlight this limitation in our modelling approach. Furthermore, we suggest running forward models for Ap-BaF-P2 to Ap-BaF-P4 using the inversion results of Ap-BaF-P1 to assess whether the cooling history derived from Ap-BaF-P1 would fit the non-invertible profiles in Apatite-BaF.  
We have revised the section and added the forward models for Ap-BaF-P2 to Ap-BaF-P4 to the appendices.
  - On the counterintuitive geometry of the  $^4\text{He}$  profiles:  
For the  $^4\text{He}$  profiles of Apatite-BaF, we observe that the  $^4\text{He}$  concentration is highest at or near the grain centre, and the lowest at the grain rims. The profiles show a concave down geometry, as expected for a grain that has experienced diffusive loss of  $^4\text{He}$ . The only exception to this pattern is Ap-BaF-P3, where the centremost spot exhibits the lowest  $^4\text{He}$  concentration (Fig. 2 and Section 3.2, lines 232-237). We did not intend to suggest that the geometry of the  $^4\text{He}$  profiles in Apatite-BaF (except for Ap-BaF-P3) is counterintuitive. We will clarify in Section 3.2 that, overall, the  $^4\text{He}$  profiles for Apatite-BaF are rather inconspicuous.  
Added clarification to Section 3.2 that profiles are inconspicuous.
  - On the central part of the grain being younger than the grain rims:  
The pattern of younger in situ AHe dates in the grain centre of Apatite-BaF compared to the grain rim is counterintuitive and is most pronounced in Ap-BaF-P2 and Ap-BaF-P3. Ap-BaF-P1, which we successfully inverted for cooling histories, does not show this trend. Thus, the overall puzzling AHe date pattern did not preclude the derivation of potential time-temperature paths for Apatite-BaF. We will revise our discussion about the spatial variation of the in situ AHe dates in Apatite-BaF for clarification, also taking into account suggestions from RC 2.  
Added discussion about He implantation, removed Apatite-URG
- **RC 2:** When the data contradict the model, it typically indicates either low-quality measurements (which I assume is not the case here) or that the model itself may not be correct. Thus, this issue should be discussed in detail. Why did the classical He-loss model fail to reproduce the results? The authors provide only a brief discussion of this matter. Among the factors considered are the presence of inclusions, uranium-enriched zones, and variations in the degree of crystallinity. However, as it was already mentioned, profiles are reproducible, no inclusions are observed in the analyzed half of the grain, and the effect of apatite crystallinity on helium diffusion is limited. The possible influence of implanted helium is not addressed.

- **AR 2:** Thank you for highlighting the issue of helium implantation. As for RC 1, we will address each aspect separately below.

- “Why did the classical He-loss model fail to reproduce the results?”

We are not entirely sure what this question alludes to. Assuming the comment targets the counterintuitive in situ AHe date pattern, for which we admittedly do not have an explanation, we suggest further expanding the discussion and elaborating on the variability in crystal lattice properties and diffusivity not directly related to the parent nuclides but to variations in major element composition (e.g. Djimbi et al. 2015) or vacancy damage (e.g. Gerin et al. 2017), which we have neglected to mention. However, with the information we have for Apatite-BaF, we cannot determine whether these factors are significant.

Added factors mentioned above to the discussion (Section 4.5)

- “The possible influence of implanted helium is not addressed.”

We thank the reviewer for pointing this out and recognise that a discussion of this topic should be added to the manuscript. We do not think that  $^4\text{He}$  implantation has a significant influence on the  $^4\text{He}$  profiles in Apatite-BaF for the following reasons:

1. When evaluating the potential impact of implanted  $^4\text{He}$  from external sources, it is commonly assumed that the outer 20  $\mu\text{m}$  rim of the apatite grain is affected the most (e.g. Spiegel et al., 2009; Gautheron et al., 2012). For our measurements, the distances of the in situ  $^4\text{He}$  ablation spots to the grain rims are larger than 20  $\mu\text{m}$ , except for two in situ  $^4\text{He}$  ablation spots that were neither included in thermal modelling nor (U-Th-Sm)/He age calculations (refer to Fig. 2 and Table 3 in the manuscript for spot distances to the grain rim).

2. Furthermore, models of implantation scenarios predict a significant peak in  $^4\text{He}$  concentration at the grain rim facing the external  $^4\text{He}$  source (see, for example, Fig. 6 in Gautheron et al. 2012). However, our Apatite-BaF  $^4\text{He}$  profiles do not show this peak, suggesting that the impact of  $^4\text{He}$  implantation is not significant.

We will add a paraphrased version of the above elaboration to the discussion section of our manuscript for clarification.

Added discussion above to Section 4.5.

- **RC 3: This point is of critical importance—if the model does not accurately describe observed in situ (U-Th-Sm)/He profiles, its validity as a basis for thermal history modeling becomes questionable.**

- **AR 3:** In our opinion, the modelling results for both Apatite-URG and Apatite-BaF clearly demonstrate that in situ  $^4\text{He}$  profiles can be inverted for plausible thermal histories (Fig. 5 and Fig. 6 in our manuscript). Given that zoned or heterogeneous grains such as Apatite-BaF are known to be problematic for conventional whole-grain (U-Th-Sm)/He approaches where rigorous criteria for optimal grain selection apply, we do not see the rationale in dismissing the in-situ  $^4\text{He}$  profile approach based on difficulties with interpreting the heterogeneous Apatite-BaF.

In our opinion, the validity of the in situ approach hinges on the ability to produce accurate results for homogeneous grains (as for whole-grain (U-Th-Sm)/He dating). We want to emphasise here that the homogeneous Apatite-URG yielded good results, and we thus consider the method valid. Apatite-BaF, on the other hand, highlights the limitations and areas of further development.

We suggest modifying our manuscript in several places to 1. explain our choice in analysing homogeneous and heterogeneous grains (testing the validity and limitations of the approach), 2. explicitly clarify that homogeneous grains are the most suitable and straightforward choice for routine analyses and 3. state that zoned or heterogeneous grains, such as Apatite-BaF, can be analysed, but are not recommended.

We added elaboration on the sample choice to Section 2.1 and added to section 3.6 that heterogeneous grains are more challenging to invert. The conclusion now explicitly states that while the method is readily applicable to homogeneous grains, heterogeneous grains are not recommended for routine analysis.

## General comments

- **RC 4:** For a study of this kind, it would have been preferable to use apatite grains from thermochronologically well-characterized localities, where FT and AHe ages are already available, or where even  $4\text{He}/3\text{He}$  profiles have been obtained. While this may not be essential from a methodological point of view, it would undoubtedly strengthen the testing and validation of the proposed thermochronological modeling approach.
- **AR 4:** This is a valid criticism, and we agree with the reviewer that a well-characterised sample or standard material would have allowed for better comparability of our approach to whole-grain (U-Th-Sm)/He dating and  $4\text{He}/3\text{He}$  analyses. Unfortunately, the well-characterised standard material available in our laboratory (Fish Canyon Tuff apatite (FCT) and Durango apatite) was not suitable for our measurement approach. For Fish Canyon Tuff, the available grains were too small. Further, zonation has been identified as a concern for this standard (e.g. Pickering et al. 2020). For Durango apatite, the grains were too large for the analysis to be performed within a reasonable cost and timeframe. For this reason, we chose Apatite-URG (from a Miocene tuff with an independently constrained U-Pb age) as a substitute for standard material. The Apatite-URG sample provided large, euhedral and clear crystals. Furthermore, the expected cooling history for Apatite-URG is simple and sufficiently well-constrained due to the volcanic nature of the source rock and the independently determined U-Pb age. Therefore, we consider this sample to be appropriate for testing and validating our methodology. We will add a few sentences explaining the choice of sample material to our manuscript. [Added explanation on choice of sample material to Section 2.1.](#)
- **RC 5:** Sample preparation involved SEM imaging and mounting in Teflon, both of which exposed the samples to elevated temperatures. The authors likely used low current values for SEM imaging and low temperatures for mounting to limit possible thermal effects; thus, it should not affect the AHe ages. But detailed information regarding this issue should be presented in the methodological section.
- **AR 5:** Thank you for pointing out the lack of discussion concerning the thermal effects of sample preparation. We will add the Teflon embedding temperature and duration (300°C for 2 minutes) as well as the SEM imaging settings (beam current 15 kV) in the methods section.

- Concerning the potential effect of Teflon embedding, we can calculate the fractional He-loss resulting from such a short-term heating event using the approximation from Reiners et al. (2007) (Equation 3). With the diffusion parameters determined for the Durango apatite standard (Farley, 2000) as an estimate and using our measured grain radii, we can approximate that the fractional He-loss from Teflon embedding is ~1% for grains the size of Apatite-BaF and ~0.5% for grains the size of Apatite-URG. We will add this information to the methods section.
- The impact of SEM analyses on the AHe system has been systematically examined by Shan et al. (2013). They found that AHe dates are not significantly affected by SEM. They also determined that the majority of heating caused by SEM would be limited to the upper 0.3  $\mu\text{m}$  of the grain. We will refer to Shan et al. (2013) in the methods section.

We added the requested details to Section 2.3 and we refer to the study by Shan et al. (2013) for the impact of SEM. The fractional He-loss calculation is added to the supplementary material.

- **RC 6: What are the reasons for the significant variations in He extraction pit depth observed within a single grain (for example, from 6.9 to 9.3  $\mu\text{m}$ ; approximately 25%; Table 3)? If the laser energy is dissipated/reflected, could it potentially heat the sample (mobilize He) or heat internal parts of the cell (increase blank level)?**

- **AR 6:** We report the maximum pit depth in Table 3. Due to the unevenness of the pit floor, the maximum pit depth can vary significantly. However, the variation in maximum pit depth does not necessarily mean the ablated volume strongly varies (Figure 1). For example, in Apatite-URG, the maximum pit depth ranges from 6.9 to 9.3  $\mu\text{m}$ , but the standard deviation of the mean pit volumes is only 4%. We suggest removing the maximum pit depth information from the manuscript and only reporting the mean pit depth.

We are unsure what the second comment regarding the dissipation of laser energy, the mobilisation of He from heating of the sample and the increase in blank level allude to in the context of pit dimension variations. As for the heating of the grain outside the pit during He extraction, Boyce et al. (2006) reported that laser heating and He degassing outside the ablation pit are negligible.

Removed the individual pit depths from Table 3 (mean pit depths are in Table 2). The deleted information is still available in supplement on Zenodo.

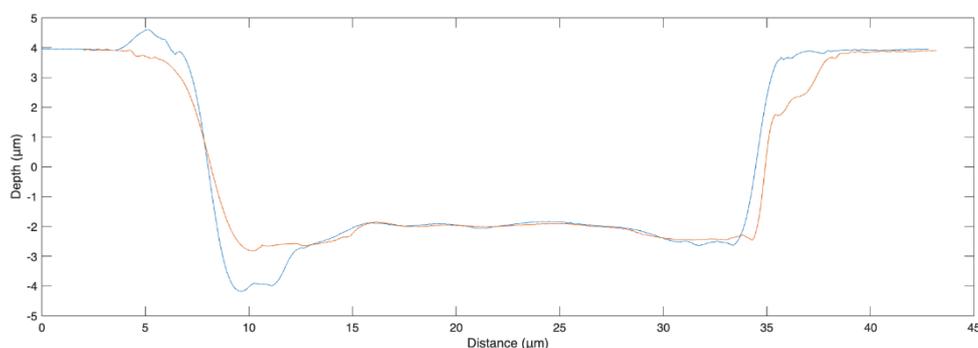


Figure 1: Cross-section through two exemplary ablation pits with significant variation in max. pit depth, but quite consistent pit volumes.

**Additional minor comments can be found within the attached file.**

### Comments from the attached pdf file

Note: The minor comments have been marked and annotated in the text and figures. We have copied the reviewer's comments and added the line information. For reference, we have also quoted text fragments marked by the reviewer and indicated which section of the manuscript the reviewer's comment concerns.

### Abstract

- Line 21 (“for our study and instrument set-up”): **provide extra details (Resochron, excimer etc)**
  - We will add “RESOchron system (Applied Spectra), consisting of a He-line and an excimer laser”.  
Added.
  
- Line 24 (“20-30microns”): **you tested only 20 microns, aren't you?**
  - In this sentence, we meant to make a recommendation for an optimal spot diameter. We will rephrase this sentence and add another sentence detailing the range of spot sizes tested (10 – 30  $\mu\text{m}$ ) and the respective grains’ AHe dates.  
Added: “We tested different laser ablation spot sizes (10-30  $\mu\text{m}$ ) in grains with a range of  $^4\text{He}$  concentrations and (U-Th-Sm)/He dates (16 to ~200 Ma)”.
  
- Line 24 (“six-spot in situ  $^4\text{He}$  profile requires a minimum grain diameter (measured perpendicular to the c-axis) of 145  $\mu\text{m}$ ”): **is it true for all apatites or only for the young ones?**
  - This is our general recommendation. With an optimal ablation spot diameter of 20  $\mu\text{m}$ , a grain size of at least 145  $\mu\text{m}$  is needed to measure a  $^4\text{He}$  profile composed of six individual in situ  $^4\text{He}$  measurements. (Details are in section 4.1 of our manuscript.) We will modify this sentence to clarify this, e.g. by writing “Additionally, **with a  $^4\text{He}$  ablation spot diameter of 20 microns**, a six-spot in situ  $^4\text{He}$  profile requires a minimum grain diameter (measured perpendicular to the c-axis) of 145  $\mu\text{m}$ ”.  
Added as suggested.
  
- Line 26 (“potentially including zoned and irregularly shaped crystals”): **so you suggest that BaF -like apatite can be used for modeling?**
  - Yes, we suggest that heterogeneous “BaF-like” apatites can be used for modelling.

### Methods

- Line 85 (“age”): **U-Pb**
  - Thank you for pointing this out. We will add U-Pb before “age standard” for clarification in line 85.  
Added.
  
- Line 90 (“embedded in a Teflon sheet”): **At what temperature?**
  - The hotplate used for embedding the grains in Teflon was set to 300°C, and the mount remained on the hotplate for 2 minutes. We will add this information.  
Added.
  
- Line 92 (“SEM”): **SEM heat the grains. What energies have you used? What was the beam current? Were the samples coated? By gold or by coal?**
  - We did not coat the sample mount. For the SEM scan, we used a voltage of 15 kV. The emission current was 40.3  $\mu\text{A}$ , the filament power was 4.46 W, and the dwell time was 200ns. We will add this information to the manuscript.  
Added.
  
- Table 1 **Which coordinate system? WGS84 or ETRS89?**
  - We use WGS84. We will add this information to the table notes.  
Added.
  
- Line 108 (“Line blanks were recorded regularly”): **It is cold blank. What about the hot blank? Have you measured the blank when ablate helium free sample?**
  - Each sample mount is individually placed in an ultra-high-vacuum (UHV) laser cell and evacuated for a minimum of 24 hours to ensure that vacuum levels reach the low-pressure regime required for high-precision He analysis. In fact, our cold cathode was unable to measure the pressure since it reached values below the detection limit  $1\text{E}-8$  torr. Before analysing samples, several initial blank analyses are run until blank  $4\text{He}/3\text{He}$  ratios reached values  $<0.0001$ . These blanks use the identical valve and volume configuration as the subsequent sample measurement, but without firing the laser.  
Added text above to supplementary material.
  
- Line 109 (“0.0003 to 0.0005 ncc”): **How do you calibrate the MS?**
  - Quantification of radiogenic  $^4\text{He}$  extracted from individual laser pits relies on calibration against an internally monitored  $^4\text{He}$  reference reservoir (“Q-tank”). The  $^4\text{He}$  volume of the Q-tank is known but not constant: it undergoes predictable exponential depletion as Q-shots are consumed during routine analytical work. To maintain accuracy, the Q-tank is recalibrated at regular intervals, typically annually or after several thousand Q-shots, using a secondary reference reservoir (“D-tank”). The D-tank contains an independently certified amount of  $^4\text{He}$ , externally calibrated by Des Patterson, and serves as the long-term stability anchor for all Q-tank recalibrations. Calibration factors derived from these intercomparisons are used to convert measured ion currents into absolute  $^4\text{He}$  amounts in sample analyses.  
Added text above to supplementary material.

- Line 111 (“after blank correction”): **have you incorporated the blank uncertainty into standard deviation calculations?**
  - Yes, blank uncertainty is propagated into the  $^4\text{He}$  concentration standard deviation. We will report the equations in the supplementary material.  
Added details on uncertainty propagation to the supplementary material.
  
- Line 122 (“re-polished grain surfaces”): **You have removed 10 microns or more?**
  - Thank you for this question. From the pit depths, we can estimate that  $\sim 10\ \mu\text{m}$  had to be polished away to smooth the grain surface. We were careful in repolishing the mount and checked the state of removal multiple times during the process to make sure we stopped polishing as soon as the pit traces had been removed. We will expand the methods section and provide more details about the repolishing procedure, e.g. by adding **“Prior to parent nuclide measurements, the grains were polished for 3.5 h on a polishing machine at intervals of 4 to 20 min with a decreasing force from 20 to 10 N to remove the helium extraction pits and create an even surface for U, Th and Sm distribution mapping. The state of removal was checked multiple times during the process, and polishing was stopped as soon as all visible pit traces had been removed.”**
  - Added bold text from above to Section 2.4.
  
- Line 124 (“Durango”): **Durango?**
  - We will add “age standard for apatite” and the reported standard (U-Th)/He age in parentheses for clarification.  
Added.
  
- Table 2 (“24”): **You have a large variations in pit depths for He measurements, and quite a constant values for trace elements. Why?**
  - Thank you for this question. We observe variations in the depths of the  $^4\text{He}$  ablation pits, but not in the depths of trace element pits, as the former are individually measured, while the latter is a calculated value. We measured the pit dimensions of each  $^4\text{He}$  ablation pit individually because the pit volume is necessary for  $^4\text{He}$  concentration calculation (lines 111 – 114 in the manuscript). The pit depth for the trace element measurements, on the other hand, is a calculated value that we derived from the established ablation-time-depth relationship (lines 129 – 130 in the manuscript). For the trace element measurements, our interest only lies in the depth in the grain from which the signal originates, not in the exact pit dimension. Hence, the calculated value suffices. We will clarify this in the table notes. We will also replace “Mean spot depth” with “Mean pit depth” in the header of Table 2 (and where relevant throughout the manuscript) and add the  $^4\text{He}$  pit depth SD for clarity.  
Changes made as suggested.

### 3. Results

- Line 213 (“of  $2.8E15 \pm 5.0E15$  at g-1 to  $8.5E15 \pm 2.5E15$  at g-1”): **why haven't you used 20 microns spot for McClure? It would allow you to receive a profile, wouldn't it?**
  - The reviewer raises a good point. We wanted to test different  $^4\text{He}$  ablation spot sizes to test the limitations of the in situ  $^4\text{He}$  measurements concerning spot size, and thus used a different ablation spot diameter for each analysed grain. We anticipated a higher  $^4\text{He}$  concentration in Apatite-McClure and therefore chose to test our smallest spot diameter in this grain. We fully agree with the reviewer that, in hindsight, a bigger spot would have been the better choice.  
Added to the discussion in Section 4.3 that ablation spot for Apatite-McClure was too small.
- Line 250 (“Apatite-McClure displays minor internal variation (...)”): **McClure apatite is well-described. Given that this data is not used in modelling I would suggest it be redundant**
  - While Apatite-McClure highlights the limitations of our method, especially concerning ablation spot size, we agree that it is of minor importance compared to Apatite-URG and Apatite-BaF. We suggest removing detailed descriptions concerning Apatite-McClure and moving the remaining information that is relevant to the manuscript to the appendices.  
We removed the sentences about Apatite-McClure and removed the parent nuclide maps of Apatite-McClure from Fig. 3. They are now displayed in the appendices.
- Line 253 (“slight (...) 5-17 ppm”): **slight? 3 times difference**
  - Both the U-concentration and the variation (5-17 ppm) are small compared to Th (86–234 ppm) and Sm (20–310 ppm). We will rephrase this sentence to clarify that the total amount and variation in U is small or “slight” relative to Th and Sm.  
Removed “slight variability” and clarified that variation in Th and Sm is greater than in U.
- Line 283 (“ $83.22 \pm 9.55$  Ma to  $162.25 \pm 28.95$  Ma”): **round the values appropriately**
  - Thank you for pointing this out, we will do that throughout the manuscript.  
Corrected in the text an in Table 3.
- Line 293 (“was 20 Ma”): **why not 16.75, or 16.75+0.84?**
  - We concede that the model box is larger than strictly necessary. We wanted to avoid the starting point dictating the inversion result. Given that the U-Pb and (U-Th-Sm)/He dates for Apatite-URG overlap, we drew a larger exploration box, including up to 20 Ma. Ultimately, this was larger than necessary, but as it does not influence the AHe date, it is a question of presentation.  
Added the sentence: “Due to the overlap in (U-Th-Sm)/He and U-Pb dates, we chose to widen the exploration space to 20 Ma instead of using the U-Pb date to avoid the starting point dictating the inversion results.”

- Line 305(“mean apatite U-Pb date derived from trace element measurements in Apatite-BaF”): **add the U-Pb data as a supplementary file**
  - The U-Pb data are already provided in the data supplement linked on Zenodo. We will refer to the supplement with a link in the text.  
Added the link.
  
- Line 306 (“near the sample location”): **where exactly? 1 km - 10 km distance? same geological unit?**
  - Thank you for this comment. Unfortunately, we can only estimate the distance of our Apatite-BaF sample location to the sampling sites of Vamvaka et al. (2014) since they do not provide coordinates, but only an overview map.  
We estimate that the distance between our Apatite-BaF sample location and the nearest sampling site of Vamvaka et al. (2014) is approximately 5-10 km. For this nearest sample, they report that the sampled rock type is granite, which aligns with the lithology of our sample. However, we cannot be certain that we have sampled the same intrusion.  
Vamvaka et al. (2014) suggest that an exhumation-and-reheating scenario is generally plausible for the Bavarian Forest. Our intention with the comparison of our Apatite-BaF to findings from Vamvaka et al. (2014) was to assess the plausibility of the derived cooling history for Apatite-BaF, not to prove its accuracy. We will emphasise this in Section 3.5.1.  
Added “Further, we used a study by Vamvaka et al. (2014) conducted near the sample location of Apatite-BaF in the Bavarian Forest (approximately 5-10 km from our sample location, same rock type), to assess the plausibility of our inversion results for Apatite-BaF.”
  
- Fig.5 and Fig. 6: **what are those artefacts near the center?**
  - Thank you for this important question. We will expand Section 2.5 to clarify the reasons for the artefacts, and we will also add a brief explanation to the figure captions of Figs 5, 6, and 7.  
The forward-modelled  $^4\text{He}$  profiles display a “jump” near the centre, because they merge two core-rim profiles. In theory, for a perfectly homogeneous grain, only one core-rim profile would be needed, since both halves of the grain (from core to rim) would yield identical information. The forward model we used (detailed in Glotzbach & Ehlers, 2024) is designed for displaying such core-rim profiles.  
To model a full rim-rim profile, which is vital for asymmetric  $^4\text{He}$  profiles (heterogeneous grains), we combine two core-rim profiles. Since both core-rim profiles share a common point at the grain centre, this centre point would be defined twice in the model. Therefore, it must be excluded (left undefined) in one of the profiles. As a result, the concentration appears to “jump” where the profiles meet.  
Added explanation to Section 2.6 and to figure captions.
  
- Line 332 as well as one apatite grain from the McClure Mountain Syenite standard (Apatite-McClure): **no  $^4\text{He}$  profiles for McClure**
  - We will remove information about Apatite-McClure from this section.

Removed

- Line 335 (“The measurement uncertainties for Apatite-McClure (ablation spot diameter 10  $\mu\text{m}$ ) exceed 40%”): **it is not a main result, i guess**
  - We will remove the information about Apatite-McClure from this section.  
Removed
- Line 346 (“sensitivity testing”): **new paragraph**
  - We will add a new paragraph.  
Added
- Line 349 (“the high individual measurement uncertainties for  $^4\text{He}$ ”): **The used ablation pit was too small. Nothing wrong with McClure -  $^4\text{He}$  concentrations are high enough. Please, rephrase**
  - We will remove the sentences referring the Apatite-McClure from this section.  
Removed. The too small ablation pit is mentioned in Section 4.3.
- Line 355 (“and heterogeneous grains”): **part of heterogeneous?**
  - We are not entirely sure what the comment intends to point out here. We suggest rephrasing the sentence for better clarity, e.g. “In situ  $^4\text{He}$  profiles can be inverted for cooling histories of homogeneous and, even though more challenging, heterogeneous grains.”  
Rephrased as suggested.

## Discussion

- Line 376 (“SD<5%”): **see table 3, SD is > 5%, while 3 blanks level is achieved**
  - We thank the reviewer for this comment. The sentence is poorly worded, and we will rephrase it.  
We ideally want a signal clearly above 3x the blank level with a standard deviation (SD) <5%. For our instrument set-up, measurements that are only slightly above 3x the blank level typically have blank-corrected SDs around 6-7%. Generally, the lower the difference between the signal and the blank, the higher the uncertainty in the measurements.  
Rephrased the sentence according to above explanation.
- Line 423 (“older dates towards the grain rim and younger dates towards the grain centre”): **It is a bit misleading, as URG might not have this trend (your next sentence)**
  - We will rephrase this section and shorten it by removing most references to Apatite-URG. We will only discuss Apatite-BaF in detail.  
Removed discussion about Apatite-URG.
- Line 429 (“In both cases, the observed date distribution within the grains is counterintuitive”): **You don't have data to proof this trend for URG sample.**

- We will remove sentences discussing Apatite-URG.  
Removed.
- Line 437 (“From our data, we cannot decipher the reason for the observed inverted in situ date pattern”): **this should be discussed in the detail! what is the most likely scenario? what about alpha implantation? whats is an average U and Th content of a host rock?**
  - We will add elaboration on the impact of  $^4\text{He}$  implantation as per our detailed response to RC 2.  
Explanation added.
- Fig. 7 **why near the zero value do you have this mathematical artefact**
  - Please refer to the above explanation for Figs. 5 and 6.  
Explanation added.
- Line 510 (“accurately”): **the error is quite large**
  - We agree with the reviewer’s point that the measurement uncertainty is relatively large. Although it may be a semantic question, it is crucial to distinguish that larger measurement uncertainties do not imply inaccuracy, but rather imprecision. The thermal histories we derive from the  $^4\text{He}$  profiles for Apatite-URG all show very rapid cooling in the Miocene, which aligns with expectations for a volcanic rock, such as the source rock of Apatite-URG. Furthermore, the timing of cooling is consistent with the independently determined U-Pb age. Thus, we maintain that the derived cooling history is accurate, even though the  $^4\text{He}$  profile measurements are arguably imprecise.

## Conclusion

- Line 558: delete “Using”
  - Ok, we can delete the word.  
Deleted.

## References

Boyce, J. W., Hodges, K. V., Olszewski, W. J., Jercinovic, M. J., Carpenter, B. D., & Reiners, P. W. (2006). Laser microprobe (U-Th)/He geochronology. *Geochimica et Cosmochimica Acta*, 70(12), 3031–3039. <https://doi.org/10.1016/j.gca.2006.03.019>

Djimbi, D. M., Gautheron, C., Roques, J., Tassan-Got, L., Gerin, C., & Simoni, E. (2015). Impact of apatite chemical composition on (U-Th)/He thermochronometry: An atomistic point of view. *Geochimica et Cosmochimica Acta*, 167, 162–176. <https://doi.org/10.1016/j.gca.2015.06.017>

Farley, K. A. (2000). Helium diffusion from apatite: General behavior as illustrated by Durango fluorapatite. *Journal of Geophysical Research: Solid Earth*, 105(B2), 2903–2914. <https://doi.org/10.1029/1999jb900348>

Gautheron, C., Tassan-Got, L., Ketcham, R. A., & Dobson, K. J. (2012). Accounting for long alpha-particle stopping distances in (U-Th-Sm)/He geochronology: 3D modeling of diffusion, zoning, implantation, and abrasion. *Geochimica et Cosmochimica Acta*, *96*, 44–56.

<https://doi.org/10.1016/j.gca.2012.08.016>

Gerin, C., Gautheron, C., Oliviero, E., Bachelet, C., Mbongo Djimbi, D., Seydoux-Guillaume, A. M., Tassan-Got, L., Sarda, P., Roques, J., & Garrido, F. (2017). Influence of vacancy damage on He diffusion in apatite, investigated at atomic to mineralogical scales. *Geochimica et Cosmochimica Acta*, *197*, 87–103. <https://doi.org/10.1016/j.gca.2016.10.018>

Glotzbach, C. and Ehlers, T. A. (2024). Interpreting cooling dates and histories from laser ablation in situ (U-Th-Sm)/He thermochronometry: a modelling perspective, *Geochronology*, *6*, 697–717, <https://doi.org/10.5194/gchron-6-697-2024>

Pickering, J., Matthews, W., Enkelmann, E., Guest, B., Sykes, C., & Koblinger, B. M. (2020). Laser ablation (U-Th-Sm)/He dating of detrital apatite. *Chemical Geology*, *548*(March), 119683.

<https://doi.org/10.1016/j.chemgeo.2020.119683>

Reiners, P. W., Thomson, S. N., McPhillips, D., Donelick, R. A., & Roering, J. J. (2007). Wildfire thermochronology and the fate and transport of apatite in hillslope and fluvial environments. 112.

<https://doi.org/10.1029/2007JF000759>

Shan, J., Min, K., & Nouri, A. (2013). Thermal effects of scanning electron microscopy on He diffusion in apatite : Implications for ( U \ Th ) / He dating. *Chemical Geology*, *345*, 113–118.

<https://doi.org/10.1016/j.chemgeo.2013.03.001>

Spiegel, C., Kohn, B., Belton, D., Berner, Z., & Gleadow, A. (2009). Apatite (U-Th-Sm)/He thermochronology of rapidly cooled samples: The effect of He implantation. *Earth and Planetary Science Letters*, *285*(1–2), 105–114. <https://doi.org/10.1016/j.epsl.2009.05.045>

Vamvaka, A., Siebel, W., Chen, F., & Rohrmüller, J. (2014). Apatite fission-track dating and low-temperature history of the Bavarian Forest (southern Bohemian Massif). *International Journal of Earth Sciences*, *103*(1), 103–119. <https://doi.org/10.1007/s00531-013-0945-x>

## Authors' response to Review 2

- Reviewer: Julien Amalberti
- Citation: <https://doi.org/10.5194/egusphere-2025-3879-RC2>

Note: The reviewer comments are presented in black print in their original, unabridged wording. [The author's response is displayed in blue.](#)

### **Review of Maier et al. (submitted to Geochronology)**

This manuscript describes a recent effort to develop in situ (U-Th-Sm)/He analytical methods and a modeling strategy applied to single apatite grains. The paper builds upon previous work by Danišik et al. (2017) and Glotzbach & Ehlers (2024), exploring specific analytical parameters such as ablation laser spot density, size, and location. The authors use natural apatite samples with both homogeneous and heterogeneous U-Th-Sm concentrations to evaluate their proposed in situ (U-Th-Sm)/He strategy. The manuscript also details the methodology used for both  $^4\text{He}$  measurement and parent nuclide analysis.

Results indicate that 4-6 laser spots, ranging from 20-30  $\mu\text{m}$  in diameter, are sufficient to derive a  $^4\text{He}$  profile from 16 Ma apatite grains of approximately 145  $\mu\text{m}$ . The authors recommend these parameters as the minimum analytical requirement for in situ (U-Th-Sm)/He analyses. They also combine  $^4\text{He}$  profiles along the c-axis (3-4 profiles per grain) with 3D reconstructions of radiogenic parent nuclide distributions to improve the accuracy of modelled thermal histories.

Overall, the paper is well written and easy to follow, but some sections lack sufficient methodological detail (see comments below). The methodology section would benefit from clearer descriptions to remove potential biases in the extracted He dataset. Moreover, the modeling approach would have been strengthened by including a case study using a well-characterized apatite standard (e.g., Fish Canyon Tuff or Durango) to test the validity of the authors' strategy.

In addition, the study decouples the He pit locations from the parent isotope analyses by repolishing the surface (of unknown thickness) between the  $^4\text{He}$  and U-Th-Sm analyses. To my knowledge, previous studies tend to perform parent isotope ablations within the same pits as  $^4\text{He}$  measurements (e.g., Pickering et al., 2020) to avoid losing spatial information during repolishing. It would have been useful if the authors provided more justification for this approach (see comments below).

Nevertheless, this study represents a meaningful advancement in in situ (U-Th-Sm)/He methodology, particularly with the introduction of the new Caw parameter for parent nuclide correction at  $^4\text{He}$  ablation sites, which accounts for surrounding  $\alpha$ -ejection effects. In my opinion, this is the most novel and valuable aspect of the study, as it tries to address  $^4\text{He}$  implementation influenced by  $\alpha$ -ejection from zones of differing eU within the grain (see

Pickering et al., 2020, Fig. 8). Therefore, I consider this manuscript a valuable contribution to the advancement of in situ (U-Th-Sm)/He thermochronology.

Below I provide my main comments, followed by detailed comments. In summary, the paper is well constructed and represents a significant contribution, but requires clarification of several methodological aspects (particularly data acquisition and error handling) and would benefit from comparison to a well-known system for validation. I therefore recommend this manuscript for publication after revision.

We thank the reviewer for their time and careful review of our manuscript. The reviewer's main concern is a lack of sufficient methodological detail. Furthermore, they express the need for a better justification of methodological choices, particularly with respect to the re-polishing step between  $^4\text{He}$  and parent nuclide measurements. Lastly, they point out the lack of analysis of a well-characterised standard material. We address these concerns in our answers to the main and detailed comments below. We are optimistic that our answers and suggested revisions to the manuscript will alleviate the reviewer's concerns.

We have numbered the reviewer comments (RC) and author responses (AR).

## Main Comments

- **RC 1: Lack of comparison with a well-known system:** Although not critical, the study would be considerably strengthened by comparison with a well-characterized apatite system such as Fish Canyon Tuff (FCT) or Durango. FCT apatite, in particular, would be an excellent candidate given recent  $^4\text{He}/^3\text{He}$  data from the distal (FCT-D) and classic (FCT-C) localities (Colleps et al., 2024). These samples provide a well-constrained framework (Renne et al., 1998; Reiners et al., 2002; Phillips & Matchan, 2013; Gleadow et al., 2015) encompassing both non-diffusive alpha-ejection-only (FCT-D) and more diffusive post-eruptive burial reheating (FCT-C) apatite behavior. Using such a reference material could effectively validate the authors' analytical and modeling strategy.
- **AR 1:** We agree with the reviewer that analysing a well-characterised standard material such as Fish Canyon Tuff (FCT) or Durango would have greatly improved the direct comparability of our approach to whole-grain (U-Th-Sm)/He and  $^4\text{He}/^3\text{He}$  analyses.

We did not analyse FCT or Durango, because the FCT grains available at our laboratory were too small for our analytical approach, while our Durango apatite material was too large to be analysed within a reasonable time and cost investment. Furthermore, FCT has been shown to exhibit significant parent nuclide zonation (e.g. Pickering et al. (2020) which disqualified it from being good validation material for our purposes.

Against this background, we chose to use the Apatite-URG sample as validation material for our in situ approach, because it contained reasonably large, euhedral, clear and homogeneous crystals. Additionally, Apatite-URG (being sourced from

Miocene foiditic tuff and with an independently determined apatite U-Pb age (Binder et al., 2023) has a sufficiently simple and well-constrained expected cooling history. We suggest modifying Section 2.1 “Samples and Sample preparation” to explicitly explain and justify our choice in sample material and address the absence of well-characterised standards with the reasoning provided above.

Added explanation to Section 2.1.

- **RC 2: Decoupling of  $^4\text{He}$  and parent nuclide locations:** The authors performed parent nuclides measurements after re-polishing the apatite surface, but the thickness removed is not specified. I am concerned that their 3D reconstructions, built from an unknown removed material layer below the  $^4\text{He}$  spots, may miss critical information. A simple addition of analyzing parent nuclides directly within or immediately below the  $^4\text{He}$  pits could have provided a valuable internal check. These data would allow the authors to test whether their calculated  $C_{\text{aw}}$  values are consistent with measured parent concentrations. Clarification and justification of this methodological choice are needed.
- **AR 2:** We thank the reviewer for this comment. We will expand Section 2.3 “Parent nuclide mapping” to include necessary details and justification for the repolishing step. Additionally, we will add annotations and a better depiction of this analytical step in Fig. 1.
  - We repolished the grain surface after the  $^4\text{He}$  measurement to map the distribution of parent nuclides across the entire exposed internal grain surface, rather than only obtaining concentrations at the  $^4\text{He}$  ablation sites. While it is theoretically possible to avoid repolishing the grain surface and instead measure the concentration of the parent nuclide in and around the  $^4\text{He}$  ablation pits, doing so in practice presents challenges. Parent nuclide signals would originate from different depths within the grain (because the measurements do not start in a flat plane). This would require us to correct for ablation depth when constructing parent nuclide maps, because parent nuclide measurements from around the  $^4\text{He}$  ablation pits would not reach the same depth in the grain as measurements conducted inside the  $^4\text{He}$  pits, introducing another source of uncertainty.
  - To polish away the visible  $^4\text{He}$  pit traces, we needed to remove a maximum of 10  $\mu\text{m}$ , based on the maximum depth of the deepest pits. We were very careful during re-polishing and checked the state of the removal regularly to avoid removing too much material. Our re-polishing protocol was as follows: “Prior to parent nuclide measurements, the grains were polished for 3.5 h on a polishing machine at intervals of 4 to 20 min with a decreasing force from 20 to 10 N to remove the helium extraction pits and create an even surface for U, Th and Sm distribution mapping. The state of removal was checked multiple times during the process, and polishing was stopped as soon as all visible pit traces had been removed.” We will add this information to the methods section.

Added the text as suggested above with minor change in wording for better integration in the manuscript. Further added: “Starting with an even surface makes spatially correlating measurements at different recorded horizontal locations in the grain easier, as the depth component can be assumed to be consistent.”

Modified Fig. 1.

- **RC 3: Methodological transparency:** Several detailed comments below relate to blanks, uncertainties, and analytical procedures. Although the authors provide substantial data, key aspects of measurement uncertainty remain unclear, particularly where modeling results appear counterintuitive. Providing additional information on blank correction and error propagation would significantly improve the reproducibility and credibility of the dataset.
- **AR 3:** We will add the requested information.  
Added details on error propagation and blank correction as supplementary material.

### Detailed Comments

- **RC 4: Line 101:** The authors mention acquisition of a  $^4\text{He}$  profile parallel to the c-axis, yet no such data are presented. Did this profile yield similar results to the perpendicular one? If so, it should be shown or discussed.
- **AR 4:** Thank you for bringing this to our attention. The data for the c-axis parallel profile are available in the supplementary material. We will refer to this in the text. We will also add sentences to the methods section explaining the reason for measuring the c-axis parallel profiles.
  - We measured a c-axis parallel profile to verify whether the  $^4\text{He}$  concentrations were consistent in the grain along this direction. In retrospect, we realised the same information can easily be obtained by comparing the perpendicular profiles, making the c-axis parallel measurements superfluous. Hence, we don't display this data in the manuscript.
  - Geometric considerations for  $^4\text{He}$  diffusion are different for c-axis parallel profiles. Thus, we did not attempt thermal modelling.  
Added “While the c-axis perpendicular profiles were acquired for thermal modelling, the c-axis parallel profiles were measured to verify the  $^4\text{He}$  concentration's consistency in the grain along this direction.” to Section 2.3.  
Added “ In the following, we describe the c-axis-perpendicular profiles used for thermal modelling. Data pertaining to the c-axis parallel profiles is not included here, as it does not provide any additional insights beyond what can be gained from comparing the c-axis-perpendicular profiles. The c-axis parallel data is available in the corresponding Zendo repository (<https://doi.org/10.5281/zenodo.15856623>).” to Section 3.2.

- **RC 5: Line 109:** 4He blanks are given in ncc, whereas Table 3 and Appendix B use atoms/g for 4He measurements. I strongly recommend using consistent units to avoid conversion errors. Reporting blanks in atoms or normalizing to ccSTP would be most appropriate.
- **AR 5:** We can convert the blanks to atoms. As atoms/g is the unit used for our inverse and forward modelling, we will continue using this unit for the tables and figures in the manuscript. The <sup>4</sup>He measurements are already reported both in ncc and in atoms/g in the supplementary data. We will refer to the supplementary material in the text.

We checked the blank level and now report it in atoms and ncc in the manuscript.

- **RC 6: Line 111:** The blank correction error for successful 4He measurements is specified as 6–15%, corresponding to apatite URG and BaF in Table 2. The blanks reported by the authors range from 0.0003 to 0.0005 cc (or  $8.07 \times 10^6$  to  $1.35 \times 10^7$  atoms), yielding an average blank of  $1.08 \times 10^7$  atoms  $\pm 2.7 \times 10^6$  atoms (or  $0.0004 \pm 0.0001$  cc). Assuming no additional error from the 4He content itself (not provided in the dataset), the propagated error based solely on the blank value would be approximately 14% for Ap-URG and 7–10% for Ap-BaF. Including the pit volume errors reported in Table 2 (~5% for URG and ~9% for BaF) increases the total 4He concentration uncertainty to roughly 15% for URG and 11–14% for BaF. These estimates, based on the average blank and its uncertainty using a weighted blank contribution factor (i.e., signal/blank ratio), are close to the authors' stated range of 6–15%, and even exceed the maximum 4He error reported for apatite BaF (<6%). This is without considering any analytical uncertainty associated with the laser ablation measurements. I therefore question how the authors derived their total error estimates. Furthermore, published RESOchron systems typically report 4He blank values between  $3.2 \times 10^7$  atoms (0.0012 ncc; Evans et al., 2015) and  $6.7 \times 10^7$  atoms (0.0025 ncc; Danišik et al., 2017), whereas the authors report significantly lower blanks ( $8 \times 10^6$  to  $1 \times 10^7$  atoms, or 0.0003–0.0005 ncc). While I acknowledge that each analytical setup is unique and background levels can vary, it would be helpful if the authors provided additional details about their analytical protocol that could explain these notably lower 4He blank values. Finally, the authors report 1 $\sigma$  uncertainties on the measured in situ AHe ages ranging from 10% to 21% for both URG and BaF, yet my own propagation of the reported errors on 4He, <sup>238</sup>U, <sup>232</sup>Th, and <sup>147</sup>Sm value (Table 3) yields roughly twice those values. It would therefore be useful if the authors detailed their analytical setup and blank correction procedure to explain these unusually low blank values and how final age uncertainties (10–21%) were propagated.
- **AR 6:** We have rechecked our blank levels and reviewed our calculations. We continue to obtain the same measurement uncertainties as reported in our manuscript.

Below, we provide details on our analytical setup and the error propagation. We will include this information as a supplementary file to our manuscript.

- **In-situ (U-Th)/He Analytical Procedure**

1. Mass Spectrometer Calibration

Quantification of radiogenic  $^4\text{He}$  extracted from individual laser pits relies on calibration against an internally monitored  $^4\text{He}$  reference reservoir (“Q-tank”). The  $^4\text{He}$  volume of the Q-tank is known but not constant: it undergoes predictable exponential depletion as Q-shots are consumed during routine analytical work. To maintain accuracy, the Q-tank is recalibrated at regular intervals, typically annually or after several thousand Q-shots, using a secondary reference reservoir (“D-tank”). The D-tank contains an independently certified amount of  $^4\text{He}$ , externally calibrated by Des Patterson, and serves as the long-term stability anchor for all Q-tank recalibrations. Calibration factors derived from these intercomparisons are used to convert measured ion currents into absolute  $^4\text{He}$  amounts in sample analyses.

2. Sample Loading and Preparation

Each sample mount is individually placed in an ultra-high-vacuum (UHV) laser cell and evacuated for a minimum of 24 hours to ensure that vacuum levels reach the low-pressure regime required for high-precision He analysis. In fact, our cold cathode was unable to measure the pressure since it reached values below the detection limit  $1\text{E}-8$  torr. Before analysing samples, several initial blank analyses are run until blank  $^4\text{He}/^3\text{He}$  ratios reached values  $<0.0001$ . These blanks use the identical valve and volume configuration as the subsequent sample measurement, but without firing the laser.

3. Measurement Sequence Design

Because Q-shots yield extremely stable  $^4\text{He}/^3\text{He}$  ratios ( $<1\%$  variability) and analytical sequences are comparatively short (typically 20–40 analyses), each sequence begins and ends with two Q-shots. No Q-shots are included between samples to avoid flushing the line with high amounts of  $^4\text{He}$  and potential memory effects. The  $^4\text{He}$  content of Q-shots is 3–4 orders of magnitude higher than the He liberated during in-situ laser extraction. For this reason, the initial Q-shots are followed by 3–4 blank measurements to allow the system to return to pre-Q-shot blank levels. Sample measurements only begin once these blank values stabilize back to the initial baseline.

Blanks are measured frequently throughout the sequence. For the first three samples, a blank is run immediately before each measurement. As the system stabilizes, the number of sample analyses between blanks is gradually increased (from one, to two, three, four, and finally up to five), while maintaining sufficient blank coverage for accurate correction.

4. Blank Characterization and Correction

For each blank and sample measurement,  $^4\text{He}/^3\text{He}$  ratios are calculated. During the Apatite-URG and Apatite-BaF analytical sessions, mean blank ratios and standard deviations were:

- 0.0000677±0.0000034 (Apatite-URG)
- 0.0000700±0.0000016 (Apatite-BaF)

With a Q-standard  $^4\text{He}$  content of 10.69 ncc for these sessions, these ratios correspond to mean blank  $^4\text{He}$  amounts of approximately:

- 0.000700±0.000046 ncc (Apatite-URG)
- 0.000757±0.000048 ncc (Apatite-BaF)

Blank correction for each sample is performed by subtracting the pre-sample blank  $^4\text{He}/^3\text{He}$  ratio from the sample's measured ratio. This ensures that sample-derived radiogenic  $^4\text{He}$  is isolated from procedural background contributions, and that time-dependent drift in blank characteristics does not introduce bias into the final (U-Th)/He age calculations.

##### 5. Helium error propagation

All measured masses are measured during 10 individual cycles (i) that last in total roughly 50 sec, and the ratio (R43) between masses 4 (M4) and 3 (M3) of each run is calculated with:

$$R43_i = \frac{M4_i}{M3_i} \quad \text{Eq. 1}$$

The mean and absolute standard deviation (SD) of the R43 measurement is calculated with:

$$\overline{R43} = \frac{1}{N} \sum_{i=1}^N R43_i \quad \text{Eq. 2}$$

$$SD_{R43} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (R43_i - \overline{R43})^2} \quad \text{Eq. 3}$$

The  $^4\text{He}$  content of Q-shot j (number of Q-shot) is derived from the depletion factor (DF = 0.999899832 with SD = 4.26343E-6) and the  $^4\text{He}$  content (11.098 ncc with SD = 0.011 ncc) determined for Q-shot 1000 with the following equation:

$$^4\text{He}_{Q_j} = ^4\text{He}_{Q_{442}} DF^{(j-442)} \quad \text{Eq. 4}$$

Accordingly, the absolute standard deviation is:

$$SD_{^4\text{He}_{Q_j}} = ^4\text{He}_{Q_i} \sqrt{\left(\frac{SD_{^4\text{He}_{Q_{442}}}}{^4\text{He}_{Q_{442}}}\right)^2 + \left((j-1000) \frac{SD_{DF}}{DF}\right)^2} \quad \text{Eq. 5}$$

To calculate the  $^4\text{He}$  content of a sample, we have to know the  $^3\text{He}$  content added to the sample. This is estimated based on the  $^4\text{He}$  of nearby Q-shots ( $^4\text{He}_{Q_j}$ ). Accounting for a drift during measuring time we fit a spline-function to  $^4\text{He}_{Q_j}$ . We used a Monte-Carlo approach to take into account the uncertainty of the  $^4\text{He}_{Q_j}$ , and randomly sampled N-times  $^4\text{He}_{Q_j}$  data from a normal distribution (using the mean  $^4\text{He}_{Q_j}$  and corresponding standard deviations). Afterwards, we evaluate the N-times fitted

functions at the measuring time of our samples and the resulting  ${}^4\text{He}_{Qt}$  and SD at measuring time are:

$${}^4\text{He}_{Qt} = \frac{1}{N} \sum_{i=1}^N f_i(t) \quad \text{Eq. 6}$$

$$SD {}^4\text{He}_{Qt} = {}^4\text{He}_{Qt} \sqrt{\frac{1}{N-1} \sum_{i=1}^N (f_i(t) - {}^4\text{He}_{Qt})^2} \quad \text{Eq. 7}$$

Blank ratios  $R_{bl}$  are calculated with Eq. 1 and used to calculate blank-correct R43 sample measurements ( $R_{Sbc}$ ) with:

$$R_{Sbc} = \frac{1}{N} \sum_{i=1}^N (R_i - R_{bl_i}) \quad \text{Eq. 8}$$

Accordingly, the absolute standard deviation is:

$$SD_{R_{Sbc}} = R_{Sbc} \sqrt{\left(\frac{SD_{RS}}{R_S}\right)^2 + \left(\frac{SD_{Rbl}}{R_{bl}}\right)^2} \quad \text{Eq. 9}$$

Where the sample and blank ratios and standard deviations have been calculated with Eq. 2 and 3. Finally, the  ${}^4\text{He}$  content of a sample at time t is calculated with:

$${}^4\text{He}_{Sbc_t} = R_{Sbc} \frac{{}^4\text{He}_{Qt}}{R_{Qt}} \quad \text{Eq. 10}$$

Accordingly, the absolute standard deviation is:

$$SD {}^4\text{He}_{Sbc_t} = {}^4\text{He}_{Sbc_t} \sqrt{\left(\frac{SD_{R_{Sbc}}}{R_{Sbc}}\right)^2 + \left(\frac{SD {}^4\text{He}_{Qt}}{{}^4\text{He}_{Qt}}\right)^2 + \left(\frac{SD_{R_{Qt}}}{R_{Qt}}\right)^2} \quad \text{Eq. 11}$$

#### 4He concentration and uncertainty:

We divide the  ${}^4\text{He}$  content ( ${}^4\text{He}_{Sbc_t}$ ) by the ablation pit volume ( $V_{Pit}$ ) to obtain the  ${}^4\text{He}$  concentration ( $C_{He}$ ). Finally, the pit volume uncertainty ( $SD_{VPit}$ ) is propagated into the total measurement uncertainty for the  ${}^4\text{He}$  concentration ( $SD_{CHe}$ ) as follows:

$$SD_{CHe} = C_{He} \sqrt{\left(\frac{SD {}^4\text{He}_{Sbc_t}}{{}^4\text{He}_{Sbc_t}}\right)^2 + \left(\frac{SD_{VPit}}{V_{Pit}}\right)^2} \quad \text{Eq. 12}$$

- **AHe age calculation and uncertainties:**

To calculate the AHe age and uncertainty we use the equations for the non-iterative solution to the age equation by Meesters & Dunai (2005).

Added all of the above as supplementary material.

- **RC 7: Line 122:** The re-polished thickness should be specified. By how much was the surface re-polished before parent isotope analysis?
- **AC 7:** We will add the requested details to the text. Please refer to **RC 2/AR 2**, where we have already elaborated in the repolishing procedure for detailed information. Now explained in methods, see AR 2
- **RC 8: Line 178:** The authors note they cannot measure ablation spots within 40  $\mu\text{m}$  of the rim due to potential  $^4\text{He}$  loss. This limitation could strongly affect the ability to model cooling histories, since rim regions often contain critical diffusional information (as in  $^4\text{He}/^3\text{He}$  methods). Have the authors evaluated the potential impact of this constraint on slow-cooled samples?
- **AR 8:** We agree with the reviewer that the inability to calculate  $C_{\text{aw}} < 40 \mu\text{m}$  from the grain rim is a major drawback of our approach.
  - We would like to note that the sentence in line 178, “*we chose not to calculate  $C_{\text{aw}}$  for  $^4\text{He}$  ablation spots with centres  $< 40 \mu\text{m}$  to the grain rim,*” refers only to  $C_{\text{aw}}$  calculation and consequently the ability to use the measurements for inverse modelling. We did not mean to state that we cannot or did not measure  $^4\text{He}$  concentrations  $< 40 \mu\text{m}$  from the grain rim. We recognise that this is worded confusingly in several places in the manuscript, and we will revise the relevant sections.
  - Although we cannot utilise the spots that are  $< 40 \mu\text{m}$  from the grain rim in the inversion for thermal histories, we can still use them for misfit calculation and comparison between measured and forward-modelled  $^4\text{He}$  profiles to ensure the inversion results are reasonable. We thus still use information from the rim region. We will emphasise this in the discussion and suggest adding annotations to both Figs. 2 and 6 to distinguish which spots were used for the inversion and which were used for forward modelling/misfit calculation.
  - For further discussion on thermal history modelling, please refer to RC 10/AR 10.  
 Added grey boxes around spots in Fig. 2 and 6 that were not used in the inversion.  
 Added explanation in method section 2.5 “This restricts the ablation spots usable for thermal history inversion to those  $> 40 \mu\text{m}$  from the grain rim. However, it does not imply that spots  $< 40 \mu\text{m}$  from the grain rim should not be measured. On the contrary, they provide crucial information about the  $^4\text{He}$  profile’s shape, and they are a vital part in assessing the quality of the inversion results through misfit calculation between the modelled and measured  $^4\text{He}$  profiles.”
- **RC 9: Lines 209-214:** I am curious where the large  $^4\text{He}$  errors ( $> 40\%$ ) originate. Based on the reported concentrations, apatite McClure appears to have  $^4\text{He}$  concentration ranging between the other two apatite grains, suggesting that each

ablation pit should provide sufficient  $^4\text{He}$  for reliable analysis. However, the laser spot size for McClure (10  $\mu\text{m}$ ) is the smallest among the samples (BaF: 20  $\mu\text{m}$ ; URG: 30  $\mu\text{m}$ ), which could potentially explain the high uncertainties if the ablated volume, and thus the total  $^4\text{He}$  signal, was too low despite appropriate concentrations.

Nevertheless, according to Appendix B (Table B1), the  $^4\text{He}$  concentrations measured in individual pits fall within the same range as those for URG and BaF. While the latter two samples display relatively low uncertainties (see also my previous comments regarding error propagation), apatite McClure exhibits unusually large errors. These discrepancies cannot be explained solely by  $^4\text{He}$  concentration or laser spot size, as the pit-by-pit  $^4\text{He}$  concentrations for McClure are intermediate between URG and BaF. In addition, the pit measurement error is reported at approximately 10% for 10 $\mu\text{m}$  laser spot size, which cannot account for the >40% uncertainties reported for McClure  $^4\text{He}$  concentrations. Could the authors clarify how these large errors were derived and identify the main contributing factors?

- **AR 9:** The large uncertainties for  $^4\text{He}$  measurements in Apatite-McClure stem from the too small ablated volume, which in turn resulted in a too low  $^4\text{He}$  signal. As the reviewer correctly noted, the  $^4\text{He}$  concentration in Apatite-McClure is comparable to Apatite-URG. However, the ablation pit volume for Apatite-McClure was only ~7% of the ablation pit volume for Apatite-URG (see Table 2 in our manuscript). Keeping in mind that the  $^4\text{He}$  concentration is the amount of  $^4\text{He}$  in a specific volume (of apatite that was ablated), it follows that if the ablated volume is small, the  $^4\text{He}$  content will also be small (assuming a constant  $^4\text{He}$  concentration).

Consequently, the  $^4\text{He}$  signal that we measured for Apatite-McClure was much lower than that for Apatite-URG, even though the concentrations are comparable. In fact, for Apatite-McClure, the  $^4\text{He}$  signal was so low that it was not significantly different from the blank level. Hence, the associated measurement uncertainty is very large. Kindly refer to AR 6 for details on error propagation. We will add a few sentences to clarify the reason for the high measurement uncertainty for Apatite-McClure in the manuscript.

Added in Section 3.1 that high uncertainty stems from too low ablated volume.

Added explanation to Section 4.3.

- **RC 10: Figure 2:** For BaF-P2 and P4, some data points lie within 40  $\mu\text{m}$  of the rim, contradicting the minimum distance limitation from the rim. How does this affect model results for these profiles (see comment line 178)? Does that be responsible for the lack of acceptable paths?
- **AR 10:** We have indeed measured  $^4\text{He}$  at a distance of less than 40  $\mu\text{m}$  from the grain rim (see AR 8).

As the reviewer points out in RC 8, excluding spots less than 40  $\mu\text{m}$  from the rims may result in the loss of crucial information. However, in terms of inverse modelling (which aims to minimise the misfit between predicted and observed data), having fewer constraints or data points to fit should make it easier for the model to find

acceptable paths rather than harder. We thus do not think having to exclude spots (e.g. for being too close to the rim for  $C_{aw}$  calculation) is the reason for the lack of acceptable paths in Ap-BaF-P2 to Ap-BaF-P4.

We observe that the un-invertible profiles Ap-BaF-P2 and Ap-BaF-P4 are significantly skewed and lack c-axis symmetry, while our inverse model is designed to operate under the assumption of c-axis symmetry. (The model is optimised for homogeneous grains.) The fact that the c-axis symmetric  $^4\text{He}$  profile Ap-BaF-P1 could be successfully inverted suggests to us that the model design and the assumption of symmetry are critical factors. We will include a few additional sentences in Section 4.7 to clarify this point.

Clarified in Section 2.5 that spots closer than  $40\ \mu\text{m}$  to the grain rims should still be measured. Rephrased section 4.8 for clarification of the inversion behaviour.

- **RC 11: Line 332:** The McClure  $^4\text{He}$  profile is not shown. Should this be included in supplementary materials?
- **AR 11:** We will remove sentences concerning Apatite-McClure from this section as they are not a main result.

Removed

- **RC 12: Line 334:** Spot size uncertainties (10-15%) appear inconsistent with volume errors reported in Table 3 (4-9%). Since pit volume depends on spot size, the propagated uncertainties should scale accordingly. Please clarify.
- **AR 12:** Thank you for bringing this to our attention. “Spot measurement uncertainties” refer to the  $^4\text{He}$  measurement uncertainties, not to He ablation pit dimensions or ablation spot size. The wording in the sentence is misleading and will be rephrased.

Rephrased. Now using “ $^4\text{He}$  measurement uncertainties”.

- **RC 13: Line 335:** The text seems to conflate  $^4\text{He}$  measurement error and spot measurement error; numbers also differ from Table 2. Please clarify what each percentage refers to.
- **AR 13:** We will re-phrase these sentences as per our response to RC 12.

Rephrased as above.

- **RC 14: Line 376:** The blank rejection criterion requires signals  $>3x$  blank according to the text. However, URG ablation pit data appear to fall below this threshold. For instance, for the Ap-URG-1 spot, the calculated  $^4\text{He}$  signal is  $2.73 \times 10^7$  atoms (using an apatite density of  $3.2 \times 10^{-12}\ \text{g}/\mu\text{m}^3$ ), whereas the average blank reported by the authors is  $1.08 \times 10^7$  atoms (or 0.0004 cc). Comparing the net  $^4\text{He}$  signal (after blank subtraction) with the blank value indicates that the measured signal is less than three times the blank. This relationship holds for all URG spots listed in Table 3, where signals are only slightly above twice the blank level. I suggest that the authors verify

these calculations and discuss how this may affect the reported  $^4\text{He}$  uncertainties for the URG apatite.

- **AR 14:** We acknowledge that this is poorly worded in the text. We will revise it. For our measurement setup, we consider a signal significantly above 3x blank level and an SD <5% as ideal. However, we do not treat this as a strict rejection criterion. The reviewer is correct in their assessment that  $^4\text{He}$  concentrations for Apatite-URG are less than 3x the blank level. We consider this still acceptable, but not ideal. Generally, the lower the  $^4\text{He}$  signal and the closer to the blank level, the larger the associated measurement uncertainties. We report the  $^4\text{He}$  measurement quality (whether the  $^4\text{He}$  signal is >3x blank level or lower) in the supplementary information. For detailed equations, please refer to AR 6.

Revised the section, paraphrased from answer above.

- **RC 15: Line 387:** What does “B1” refer to in Table 2?
- **AR 15:** This should read Table 2 and Table B1. We will fix this.  
Fixed.
- **RC 16: Lines 422 - section 4.4:** The authors report an interesting and somewhat unexpected result in their modelled ages: the crystal cores yield younger ages, while the rims show older ages. This pattern is indeed counterintuitive. The authors suggest that this discrepancy might stem from the material removed during re-polishing between the  $^4\text{He}$  and parent nuclide analyses. However, it is unclear why they did not acquire additional data from the same layer as the  $^4\text{He}$  measurements, either directly within the initial laser spots (as done by Pickering et al., 2020) or from nearby locations. Such data could have provided a valuable cross-check to support the modelled Caw values. Because the thickness of the polished layer is not specified, it is possible that too much material was removed, potentially invalidating the 3D extrapolation at the actual  $^4\text{He}$  spot depth. A quick verification at or near the  $^4\text{He}$  ablation sites would have allowed the authors to confirm or adjust their model accordingly. As it stands, the observed age pattern is difficult to interpret and may indicate underlying issues with either the dataset or the modeling approach.
- **AR 16:** We kindly refer to RC 2/ AR 2 for the reasoning behind the re-polishing and the amount of material removed. We are unsure how measuring “directly within the initial laser spots (as done by Pickering et al., 2020)” as the reviewer suggests, would yield better information compared to our approach. Measuring “in” the ablation spot also means accessing information from below the  $^4\text{He}$  ablation layer. This is nicely illustrated in Fig. 4 of Pickering et al. (2020). In our opinion, we do essentially retrieve information from the same layer as we would when ablating inside the  $^4\text{He}$  pits.  
Further, in section 4.4, we initially aimed to highlight that the grinding and polishing required to expose the internal grain surface after Teflon mounting result in the loss of approximately half of the grain and thus lead to a significant loss of information for

heterogeneous grains. We will clarify in section 4.4 which specific polishing step we are referring to.

Details for the repolishing step added as per RC2/AR2.

Added clarification that we refer to “the portion of the grain that was lost during the initial grinding and polishing of the grain after it was embedded in Teflon”.

- **RC 17: Line 482-483:** The comparison presented for the BaF model (Fig. 7b) would have been more informative if the modelled curves had included a zoned parent nuclide distribution rather than only a homogeneous system. As the authors correctly point out, a homogeneous model does not adequately represent this grain’s characteristics. Incorporating a modelled zoned profile based on the observed 2D parent nuclide distribution would have provided a more meaningful test of whether the Caw correction significantly improves the model outcomes, since its benefit is not particularly evident in the homogeneous case.

- **AR 17:** Thank you for this good suggestion. We will add such a profile to Fig. 7b. We have replaced the homogeneous model with the suggested heterogeneous profile.

- **RC 18: Figure 7:** Why is the 5x5  $\mu\text{m}$  grid model for URG not shown? Additionally, since increasing grid resolution seems to improve model outputs, it would be helpful to test finer resolutions (e.g., 1  $\mu\text{m}$ ) and discuss limiting factors.

**AR 18:** We do not show a 5x5  $\mu\text{m}$  grid model for Apatite-URG in Fig. 7, because we did not calculate one. We agree with the reviewer that a higher grid resolution appears to improve the model outputs. However, as Fig. 7a shows, the 10x10  $\mu\text{m}$  grid model is already indistinguishable from the homogenous model within measurement uncertainty, and the misfits between modelled and measured  $^4\text{He}$  profiles are nearly the same. Hence, we do not see justification for generating even finer resolution maps for Apatite-URG. We will add this reasoning for the missing 5x5  $\mu\text{m}$  grid model for Apatite-URG to the manuscript.

Regarding the general testing of finer interpolation resolutions and their limiting factors, in our opinion, this should be addressed in a separate study. It is important to not only systematically test the interpolation resolution for parent nuclide maps, but also the combination of ablation spot size and interpolation grid resolution. This is beyond the scope of the manuscript.

Added explanation to Section 4.7. “Given that both models are indistinguishable within measurement uncertainty, we did not generate finer resolution models”.

Rephrased sentence for clarification that map resolution should be explored in combination with ablation spot size: “Further studies are also required to test the optimal interpolation grid resolution in combination with the ablation spot size and the necessity of element maps of the entire grain.”

- **RC 19: Line 510-512:** The authors claim that six laser spots are sufficient to reconstruct a simple, rapid cooling history with minimal or no parent nuclide zonation. However, for more complex and strongly zoned samples, it appears that

only five laser spots (e.g., Ap-BaF-P1) were enough to produce acceptable time–temperature paths. I find this conclusion somewhat surprising and possibly overinterpreted. Could the authors clarify how five laser spots were sufficient in the more complex case, and whether this is supported by model sensitivity tests?

- **AR 19:** We will modify the text to clarify that the minimum number of spots needed for a  $^4\text{He}$  profile is not strictly implied by our results. We will also add a paraphrased version of our explanation below to the manuscript.

Generally, we recommend using as many points as possible to measure a  $^4\text{He}$  profile. As the reviewer points out, we cannot derive the minimum number of spots in a  $^4\text{He}$  profile purely based on our measurements. The number that we give is our best estimate based on the data that we have available and a mathematical thought experiment: The minimum number of  $^4\text{He}$  spots required to define a  $^4\text{He}$  profile can be compared to determining how many points are needed to define a curve. It is trivial that two points define a straight line, and given any two distinct points, there is only one unique line that can be drawn through them. The minimum number of points required to define a curve is a more nuanced problem, as it depends on the complexity of the curve. The simplest assumption is that if the definition of a straight line requires two points, a simple curve should require at least three points. Indeed, it can be shown that a simple quadratic function of the form  $f(x) = ax^2 + bx + c$  is uniquely determined by three distinct points.

If we assume, in the simplest mathematical case, that the  $^4\text{He}$  profile within a homogeneous grain has a parabolic shape, we would thus only need three points from the core to the rim to measure the  $^4\text{He}$  profile. (We cannot make the same assumption for rim-to-rim profiles, because three points would mean defining one shared central point and one outer point for each half of the grain. This would result in describing a line in each grain half, rather than capturing a curved profile.)

For rim-rim profiles, we would double the number required for a core-rim profile, resulting in six spots. In the case that the  $^4\text{He}$  profile is parabolic in shape and symmetric about the c-axis (meaning the vertex of the parabola is located at the grain centre), five spots can be sufficient. Then, the middle spot should be located exactly at the grain centre. By chance, Ap-BaF-P1 is sufficiently close to that condition, and five spots are enough for measuring the  $^4\text{He}$  profile.

We do not know beforehand how complex the  $^4\text{He}$  profile's shape in the grain will be, and thus how many measurements are needed. It can very well be more than six. However, we can conclude with some confidence that one should not start with less than six (rim-rim).

Added a new paragraph with the explanation from above (slight modified).

- **RC 20: Line 519-521:** Do the blue, red, and black curves shown in Fig. 7 represent two separate models extending from the core to the rim on each side (i.e., toward positive and negative distances from the center)? If so, this could explain the slight discrepancies observed between the curves at the grain center, where the two models meet. This point is not enough clearly stated in the text to my opinion, and it would be

important to clearly specify that these are half-grain models (core-to-rim) rather than full-grain simulations (rim to rim). Is the same modeling approach also applied to the URG apatite? Similar minor mismatches are visible near the center of those plots as well.

- **AR 20:** We indeed merge two core-rim profiles for forward-modelling. We will add explanations to the text and modify the figure captions to highlight this.  
Explanation added to Section 2.6 and referred to in figure captions.
- **RC 21: Appendix B:** This table provides information for the McClure sample but does not include the calculated in situ AHe date. Although this sample was not used in the modeling section, the authors successfully measured  $^4\text{He}$  signals and corresponding Caw U-Th-Sm values for numerous ablation pits. Do these individual ages combine to yield a whole-grain equivalent age given at  $\sim 523$  Ma (see table 1)? it would be useful to include calculated in situ AHe dates, as it provides additional context and comparability with URG, especially given similar parent nuclide homogeneity (Fig. 3).
- **AR 21:** We did not calculate AHe dates because we were doubtful of their significance, given the large uncertainties on the  $^4\text{He}$  concentrations. It is an interesting question how the in situ AHe dates compare to the crystallisation age reported for the McClure Mountain Syenite. Previous studies (e.g., Anderson et al., 2017) report AHe dates that are significantly younger than the crystallisation age of  $\sim 523$  Ma, ranging from  $\sim 100$  to 140 Ma. We will add AHe dates for Apatite-McClure to the Appendix.  
Added the AHe dates.

## References

Danišik, M., McInnes, B. I. A., Kirkland, C. L., McDonald, B. J., Evans, N. J., and Becker, T. 2017: Seeing is believing: Visualization of He distribution in zircon and implications for thermal history reconstruction on single crystals, *Science Advances*, 3(2), 1–10,

<https://doi.org/10.1126/sciadv.1601121>

N. Evans , J. Byrne , J. Keegan and L. Dotter LE , Determination of Uranium and Thorium in zircon, apatite, and fluorite: application to laser (U–Th)/He thermochronology, *J. Anal. Chem.*, 2005, 10.1007/s10809-005-0260-1, 60, 1159 -1165.

Glotzbach, C. and Ehlers, T. A.: Interpreting cooling dates and histories from laser ablation in situ (U–Th–Sm)/He thermochronometry: a modeling perspective, *Geochronology*, 6, 697–717, <https://doi.org/10.5194/gchron-6-6972024> , 2024.

Pickering, J., Matthews, W., Enkelmann, E., Guest, B., Sykes, C., and Koblinger, B. M., 2020, Laser ablation (U-Th-Sm)/He dating of detrital apatite: *Chemical Geology*, v. 548, p. 119683. <https://doi.org/10.1016/j.chemgeo.2020.119683>

Phillips, D., and Matchan, E., 2013, Ultra-high precision  $^{40}\text{Ar}/^{39}\text{Ar}$  ages for Fish Canyon Tuff and Alder Creek Rhyolite sanidine: new dating standards required?: *Geochimica et Cosmochimica Acta*, v. 121, p. 229-239. <https://doi.org/10.1016/j.gca.2013.07.003>

Reiners, P. W., Farley, K. A., and Hickey, H. J., 2002, He diffusion and (U-Th)/He thermochronometry of zircon: initial results from Fish Canyon Tuff and Gold Butte: *Tectonophysics*, v. 349, no. 1-4, p. 297- 308. [https://doi.org/10.1016/S0040-1951\(02\)00058-6](https://doi.org/10.1016/S0040-1951(02)00058-6)

Renne, P. R., Swisher, C. C., Deino, A. L., Karner, D. B., Owens, T. L., and DePaolo, D. J., 1998, Intercalibration of standards, absolute ages and uncertainties in  $^{40}\text{Ar}/^{39}\text{Ar}$  dating: *Chemical Geology*, v. 145, no. 1-2, p. 117-152. [https://doi.org/10.1016/S0009-2541\(97\)00159-9](https://doi.org/10.1016/S0009-2541(97)00159-9)

Additional references cited by the authors in their response:

Anderson, A. J., Hodges, K. V., & van Soest, M. C. (2017). Empirical constraints on the effects of radiation damage on helium diffusion in zircon. *Geochimica et Cosmochimica Acta*, 218, 308–322. <https://doi.org/10.1016/j.gca.2017.09.006>

Binder, T., Marks, M. A. W., Gerdes, A., Walter, B. F., Grimmer, J., Beranoaguirre, A., Wenzel, T., & Markl, G. (2023). Two distinct age groups of melilitites, foidites, and basanites from the southern Central European Volcanic Province reflect lithospheric heterogeneity. *International Journal of Earth Sciences*, 112(3), 881–905. <https://doi.org/10.1007/s00531-022-02278-y>

Meesters, A. G. C. A., and T. J. Dunai (2005), A noniterative solution of the (U-Th)/He age equation, *Geochem. Geophys. Geosyst.*, 6, Q04002, doi:[10.1029/2004GC000834](https://doi.org/10.1029/2004GC000834).