Response letter – Anonymous Referee #1

Comments of reviewers in black, author responses in blue

This paper presents a large number (n > 150) of MC-ICP-MS U-series dating of stalagmites from a cave in the northeastern Yucatán Peninsula, Mexico. These stalagmites are particularly important because they cover the entire period of Maya cultural evolution and can be used to reconstruct a continuous, high-resolution climate record over the past 2700 years. These samples are characterized by high U and low Th concentrations, which is generally favorable for precise U-series dating. However, dating results reveal significant age inversions. These inversions are attributed to elevated initial 230Th/232Th ratio, ranging from 4 to 68, as estimated through multiple approaches, including U-series dating of modern carbonate and drip water, isochron analyses, stratigraphic constraints, and lamina counting. This paper demonstrates the potential of large variations of initial 230Th/232Th ratio in speleothem, which may pose a substantial challenge for accurate U-series dating. I think this is well-organized study supported by convincing evidence.

We thank referee #1 for their positive feedback to our manuscript.

My major comments are given below:

- The study focuses on stalagmites from Áaktun Kóopo Cave, located in the northeastern Yucatán Peninsula, and cites several other stalagmites records from elsewhere on the Peninsula. Is the elevated initial 230Th/232Th ratio specific for Áaktun Kóopo Cave, or does it reflect a regional characteristic?

We thank Referee #1 for this important question. Elevated and highly variable R02i, like those observed in Áaktun Kóopo’ Cave, are not unique to this site. Previous studies from the Yucatán Peninsula have already reported similarly high initial ratios based on isochrons or stratigraphic constraints, drip water and zero-age carbonate (e.g., Ridley et al. (2015); Moseley et al. (2015); Kennett et al. (2022); Stinnesbeck et al. (2020)). However, systematic, cave-by-cave surveys are still sparse, so it is not yet possible to map this behavior at the scale of the entire peninsula. Nevertheless, the existing evidence strongly suggests that elevated and temporally variable R02i, is a regional characteristic of many (sub-) tropical karst systems, rather than a peculiarity of Áaktun Kóopo’ alone. We now highlight this point more clearly in the revised conclusions by adding the sentence:

“Considering the vast number of other studies suggesting high and variable R02i ratios in (sub-) tropical caves, we expect that this could be a characteristic of other, comparable cave systems.” (L572)

In addition, we will change the statement in L73 of the introduction to:

“*These observations highlight the need for more systematic assessments of initial Th variability when correcting ^{230}Th/U ages from such environments*.”

- Samples with (^{230}Th/^{232}Th) activity ratio <200 were excluded from age modeling, and their corrected ages are not presented in Table S3. However, their initial (^{230}Th/^{232}Th) activity ratios are still presented. I suggest either presenting both the corrected ages and the initial values, or removing both sets of values. In the main text, the initial (^{230}Th/^{232}Th) activity ratios are reported between 4 and 68, but the corrected ages corresponding the higher initial values (>50) were discarded. Does this imply that the derived high initial values (>50) are unreliable?

We thank Referee #1 for this careful observation. Samples with ($^{230}Th/^{232}Th$) < 200 were excluded from age-depth modelling because the age uncertainty for these highly Th-contaminated samples is dominated by the poorly constrained initial Th correction, so their corrected ages would not add meaningful information to the chronologies. For this reason, we do not report their corrected ages in Table S3.

We nevertheless report the corresponding initial ($^{230}Th/^{232}Th$) activity ratios for all dated samples, including those excluded from modelling, because these ratios themselves help to characterize the range and variability of initial Th in the system. In other words, for low ($^{230}Th/^{232}Th$) samples the corrected ages are not considered robust, but the inferred R02i values still provide qualitative evidence that very high initial ratios can occur.

Regarding the highest initial ratios (>50), these mostly arise in strongly Th-contaminated samples with large, propagated age uncertainties, so the exact numerical values are not tightly constrained and should be interpreted with caution. However, the fact that independent approaches (isochrons, stratigraphic constraints, and layer-counted chronologies) all require elevated initial ($^{230}Th/^{232}Th$) to restore stratigraphic order indicates that high initial values are plausible and likely necessary. To avoid misunderstanding, we will clarify in the Methods/Supplement that (i) samples with ($^{230}Th/^{232}Th$) < 200 are excluded because their ages do not add value to the age models, and (ii) the reported initial ratios for these samples are used only to illustrate the range of initial Th contamination, not as precise constraints on individual ages. In detail, we will change the last paragraph in section 4.1 to

"Given the difficulty in accurately constraining R02i, all samples with ($^{230}Th/^{232}Th$) ratios <200 (n=68) were excluded from the age–depth modelling and from further chronological interpretation (Figure 2b), because for such highly Th-contaminated samples the propagated age uncertainty is dominated by the initial Th correction and does not add meaningful information to the chronologies. This criterion also applies to all samples within 7.5 mm of the stalagmite tops, which are generally highly contaminated and have limited ^{230}Th ingrowth."

In addition we will add the following statement to the end of 4.1.1:

"Taken together, these independent approaches indicate that strongly elevated initial Th activity ratios are plausible and necessary at least in parts of the record to obtain stratigraphically consistent age models, even though the exact numerical values of R02i for individual, highly contaminated samples remain poorly constrained."

In addition, we will add a clarifying statement in the caption of Table S3:

"For samples with ($^{230}Th/^{232}Th$) < 200, corrected ages are not reported, because age uncertainties are dominated by initial Th correction and do not constrain the chronologies. The corresponding initial ($^{230}Th/^{232}Th$) ratios are shown only to illustrate the range of initial Th contamination and should not be treated to calculate ages for individual samples."

- The sample EO-C-81.5 has the lowest Th concentration (~0.00013 ppb) and the highest (230Th/232Th) activity ratios (213749), yet its corrected age is still not consistent with the stratigraphic order.

Thank you for spotting this. Indeed, the measurement E0-C-81.5 constitutes an outlier within the E0-C dataset: despite its very low ^{232}Th concentration and extremely high (^{230}Th/^{232}Th) activity ratio, the corrected age is not consistent with the surrounding stratigraphy. In line with our general strategy, this analysis was therefore excluded from the age–depth modelling and from any chronological interpretation. We retain the measurement only in the full analytical table for completeness, but do not use it to constrain the chronology or to infer initial Th systematics.

- In Table S3, there are two rows of data for sample E6-31.5 and three rows of data for E6-220. Are these replicated analyses? Given the challenging of the Áaktun Kóopo Cave samples, it would be beneficial to present replicate analyses to show the reproducibility.

Thank you for spotting this. The entries for stalagmite E6 in the original Table S3 mixed older and more recent measurements that had not yet been evaluated consistently with the rest of the dataset. We have now corrected Table S3 so that only the final, consistently processed results for E6-31.5 and E6-220 are reported, and we apologize for this oversight.

Regarding replicates: In fact, we tried repeat analyses in several key intervals, in particular within stalagmite E0-C. For example, we measured three samples at 75–77 mm depth, and in this case the resulting data agree within their analytical uncertainties. However, it was not always possible to obtain sufficient material at exactly the same depth (e.g., where growth rates are slow or variable), so replicated ages are only available for selected horizons rather than systematically throughout all stalagmites.

Response letter – Referee #2 (Philippe Claeys)

Comments of reviewers in black, author responses in blue

This manuscript addresses the challenge of uncertainty in U-Th dating due to high, complex initial thorium contamination in stalagmites. By integrating isochron analysis, stratigraphic constraints, and annual layer counting to determine the initial $^{230}Th/^{232}Th$ ratio successfully, and by establishing multiple age models to construct robust chronologies over the past 2.7 kyr in the Áaktun Kóopo cave in Yucatan, it provides a crucial chronological foundation for studying hydrological and environmental changes during the Mayan social upheaval. The study indicates that assuming a fixed initial thorium concentration may lead to significant dating biases, with important methodological implications for high-precision stalagmite chronology research.

Dear Philippe Claeys, we thank for the positive feedback to our manuscript.

1. Based on the titles and content, this manuscript is a methodological study to constrain initial thorium concentrations through several approaches and to determine a robust speleothem Age-depth model, which is then used to reconstruct hydrological change in the Yucatán Peninsula, Mexico. The emphasis should be placed more strongly on the process of method development, with the hydroclimatic reconstruction serving merely as an illustrative application. Currently, the organization/content of the abstract appears a bit unbalanced, with a relatively greater focus on presenting the paleoclimate results compared to the methodology used.

We thank for this suggestions, and we agree with the reviewer. We will revise the abstract accordingly to put the emphasis more on the methodology:

“Speleothems, such as stalagmites and flowstones, are invaluable archives of past environmental and climatic conditions due to their layered growth and suitability for precise 230Th/U dating. In tropical karst settings, however, elevated and variable detrital thorium contamination can severely compromise age accuracy, especially for young speleothems. Here, a suite of stalagmites from Áaktun Kóopo’ Cave from the Yucatán Peninsula is used as a case study to develop and evaluate a multi method framework for constraining initial thorium and constructing robust age–depth models. We combine extensive 230Th/U dating with local isochron analysis, stratigraphic approaches, and annual Sr/Ca layer counting to constrain elevated and initial (230Th/232Th) activity ratios and resolve pronounced chronological inversions. High uranium concentrations in the speleothems (average c. 1ppm) allow precise measurements, yet we infer a high and largely unsystematic variability of high initial (230Th/232Th) activity ratios in space and time, with values spanning between 4 and 68, thus far exceeding standard bulk Earth values. This variability demonstrates that elevated and temporally variable initial Th may be common in tropical karst settings and that multiple, independent constraints on initial (230Th/232Th) activity ratios are essential for reliable chronologies in such environments. Applying this framework to Áaktun Kóopo’ cave stalagmites yields internally consistent stalagmite age–depth models spanning the past 2.7 kyr, as well as evidence for earlier glacial and interglacial growth phases. These chronologies provide the basis for high resolution, multi proxy reconstructions of hydroclimate and vegetation change in the northeastern Yucatán Peninsula during the entire era of Maya cultural evolution, and they illustrate how robust speleothem chronologies can be obtained even in cave systems affected by elevated and highly variable initial thorium.”

Furthermore, although the introduction highlights the importance of constraining initial thorium for U-Th dating through a progressive line of reasoning, it should more clearly address the specific challenges posed by initial Th uncertainty in U-Th dating, as well as the limitations of the existing approaches to correct for initial thorium.

We will revise the introduction accordingly. For example, we will adjust L73ff to:

"These observations thus demonstrate a vast range of initial Th variability, also within single cave sites. However, large (and poorly constrained) variations in R02i directly translate into additional, often dominant, age uncertainties, particularly for young and/or strongly contaminated speleothems. Classical correction schemes based on a single bulk Earth detrital ratio can therefore produce stratigraphic age inversions and biased chronologies when applied in such settings."

In addition, we will change L87ff to:

"*While monitoring detrital activity ratios in ZAC and drip waters provides useful information about initial activity ratios, it offers little insight into whether and how these ratios change over time. Overall, most studies still rely on only one of these methods, and the resulting uncertainty in initial Th variability is often poorly quantified. This highlights the need for more systematic assessments of initial Th when correcting 230Th/U ages from such environments*."

2. References about the Isochron method should be added (line 78). This is important.

We will add relevant references, e.g., (Kaufman, 1993; Richards and Dorale, 2003; Frank et al., 2002; Ludwig and Titterington, 1994)

3. In the introduction or study site part, there should be more details about the climatic, geological, and regional hydrological conditions of the study area. The initial ($^{230}Th/^{232}Th$) activity ratio of detrital material can vary with the characteristics of the host rock, soil, and chemical complexing agents during the aqueous transport of Th. Thus, it is better to give a clear description of the climatic, geological, and regional hydrological background.

We will add the additional information.

For example, we will add the climatological information at the beginning of section 2.1:
"This region experiences a warm, humid tropical climate with highly seasonal rainfall (wet season June–November), high year round humidity and frequent tropical storms and hurricanes (Romero and Alfaro, 2024; Medina-Elizalde et al., 2016a) and the soils are typically thin and variable, with pockets rarely reaching 0.5 m between areas of exposed bedrock (Lases-Hernandez et al., 2019)." (L115ff)

In addition, we will specify the geological setting as:

"Áaktun Kóopo' Cave is developed within the low relief Yucatán carbonate platform, which consists of thick Cretaceous and Tertiary shallow marine limestones, forming a great limestone plain with a shallow water table and almost no surface rivers (Ward et al., 1995; Perry et al., 2009; Smart et al., 2006)." (L120ff)

4. The relative height or elevation information of the cave samples should be discussed because flood events can disrupt the stability of cave environments through clay input and sedimentary discontinuities.

We will add an according statement in section 4.2.2:

"*One prominent flooding layer (~ 1 cm thick) appears in stalagmites E0-C, E1, and E23-3, which were all collected from the same cave chamber at the same elevation. This layer likely represents* […]".

As well as

"*Thinner layers (0.5–1 mm) in stalagmites E1 and E8, which also stems from the same area in the cave, correspond to a shorter interruption […]".*

5. The clarity and quality of Figure 1 both need to be significantly improved. Also, Figure 1b seems unnecessary, since this information is already shown in Figure 1a. It is recommended to provide more detailed information on the cave environment itself, and, in particular, the places where the walls were modified. see P. 5

We will revise Figure 1, and, in particular, remove Fig. 1b. In addition, we will highlight the information about the location of Mayan walls in the caption.

6. Methodology: Since this manuscript addresses the challenges of the initial Th, it is recommended to add more details about the U and Th extraction processes used in this study. see P. 8 205-215.

We will add more details about the protocol and extraction process in the according section, e.g.,

"For the analysis, small sample plates weighing approximately 100 mg were cut perpendicular to their growth direction using a diamond wire saw and subsequently sample leaching in weak acid, weighing, dissolution in ultra-clean nitric acid and spiking with our in-house triple-spike solution ('TriSpike', 233U, 236U and 229Th) (Kerber et al., 2023). For drip water analysis, about 15 to 30ml of sample was weighed and spiked with 'TriSpike', acidified with ultra-clean nitric acid, subsequently dried down and re-dissolved. Overall, the preparation involving sample leaching, dissolution, spiking, and wet column extraction chromatography using Eichrom UTEVA resin, followed established protocols (Wefing et al., 2017; Matos et al., 2015; Douville et al., 2010). Chemical blanks are regularly monitored and commonly below 0.4 fg for 234U and 0.04 fg for 230Th (Kerber et al., 2023; Kerber et al., 2025).Chemical blanks are regularly monitored and commonly below 0.4 fg for ^{234}U and 0.04 fg for ^{230}Th (Kerber et al., 2023; Kerber et al., 2025)." (L218ff)

Also, the methodology used to measure drip water is not provided; see Table S2.

Drip water and speleothem samples were analyzed with the same methodology. We will add the according information:

"*A total of 157 samples from Áaktun Kóopo' speleothems as well as the three drip water samples were analyzed using high-precision ^{23}0Th/U dating performed with a multi-collector inductively coupled plasma source mass spectrometer (MC-ICP-MS, Thermo Fisher Neptune plus) at the Institute of Environmental Physics, Heidelberg University. […] For drip water analysis, about 15 to 30ml of sample was weighed and spiked with 'TriSpike', acidified with ultra-clean nitric acid, and subsequently dried down and re-dissolved."*

Specific half-lives should also be provided, rather than just citing Cheng et al. (2013).

We will add the following more specific statement:

"Ages were calculated using the half-lives provided by Cheng et al. (2013) for ^{230}Th and ^{234}U as well as by Jaffey et al. (1971) for ^{238}U." (L230ff)

Questions:

- This ms focuses on the impact of high detrital Th on U-Th dating errors and emphasizes the importance of initial Th correction. However, it attributes anomalous fluctuations in high detrital Th concentrations in tropical regions as the rule, without delving into the controlling mechanisms. Does this fundamentally undermine the logical basis of the initial Th correction? In other words, if the specific processes and driving factors behind anomalous Th signals cannot be elucidated, will the effectiveness and applicability of isochronous ages calculated based on the assumptions of "uniform initial ratios" or "statistical averages" necessarily face significant uncertainties?

We thank Referee #2 for raising this important conceptual point. We fully agree that the processes and sources controlling detrital Th transport and initial ^{230}Th/^{232}Th in tropical karst systems are complex and not yet well constrained. This is precisely why assuming a single, fixed "bulk Earth" initial ratio can be problematic and may bias ages, especially in young, highly contaminated speleothems.

This complexity does not, however, invalidate the logic of applying initial Th corrections in general. Rather, it means that such corrections must be treated as approximations with explicit uncertainty, and that they should be constrained as far as possible by independent information. In our study, we therefore combine and compare several approaches to estimate R02i.

Of course, in an ideal case one would work only with material that has very low detrital Th and minimal initial-ratio complexity. In Áaktun Kóopo and similar sites of high archaeological and paleoclimate interest, such "ideal" speleothems are not always available, so dealing with significantly Th-contaminated material is unavoidable if one wishes to obtain any chronological information at all. In this context, independent constraints such as layer counting do not make Th corrections obsolete, but they can reduce the degrees of freedom in choosing R02i and provide a powerful test of whether a given correction scheme yields realistic ages. Our main conclusion is therefore not that anomalous Th behavior removes the usefulness of initial Th corrections, but that it necessitates (i) moving beyond a single fixed initial ratio, (ii) using multiple, independent constraints whenever possible, and (iii) explicitly propagating the remaining uncertainty into age–depth models.

To make this reasoning also clearer in the revised manuscript, we will add a statement to section 4.1.2:

"In such settings, initial Th corrections are best treated as approximations with explicit uncertainty, constrained where possible by multiple, independent estimates of R02i." (L395ff)

As well as to the conclusions:

"While low-Th speleothems with simple initial ratios would be preferable, such material is not always available at sites of high archaeological and palaeoclimate interest, so working with variable Th-contaminated samples is sometimes unavoidable if any chronological information is to be obtained." (L555ff)

- The validity of the isochron approach fundamentally relies on homogeneous initial thorium isotopic compositions among co-genetic subsamples. However, the manuscript explicitly reports highly variable and largely unsystematic initial ($^{230}Th/^{232}Th$) activity ratios. How can the authors justify the assumption of isotopic homogeneity required for meaningful isochron construction under such strongly heterogeneous conditions?

We thank Referee #2 for discussing this important point. We fully agree that the classical isochron approach is formally valid only when co-genetic subsamples share a common initial Th isotopic composition. In Áaktun Kóopo', we therefore use isochrons only as local, first-order approximations applied to closely spaced subsamples within limited depth intervals, where the assumption of similar initial Th is most likely to hold. The resulting R02i values are treated as approximate rather than exact and are not assumed to represent the entire cave system.

In practice, these local isochron estimates are used alongside other approaches (e.g. stratigraphic reasoning and, where available, Sr/Ca layer counting) to narrow the plausible range of R02i and to quantify the associated age uncertainty. Our intention is not to claim perfectly homogeneous initial Th, but to show that, even under heterogeneous conditions, the isochron approach can still provide useful first-order constraints on R02i when interpreted cautiously and in combination with independent information.

To acknowledge this simplification, we will change the first paragraph of section 4.1.1 to

"*Dense sampling in some stalagmites allows for first-order linear two-endmember mixing regressions (Osmond type 1- like 'isochrons"). In this study, such 'isochrons' are applied only to closely spaced, co-genetic subsamples and are interpreted as local, first-order estimates of R02i, similar to the approaches of Warken et al. (2020) or Stinnesbeck et al. (2020). The 'isochron' approach using ISOPLOT (Ludwig and Titterington, 1994; Ludwig, 2008; Vermeesch, 2018), produced five successful 'isochrons' from three stalagmites (E0-C, E4, E8), which yielded R02i ranging from 6.4 ± 2.4 to as large as 19 ± 10." (L353ff)*

In addition, we will modify the conclusion section to:

"To determine R02i and resolve chronological inversions, we successfully employed a multi-method approach, combining isochron methods, stratigraphic constraints, and relative seasonal layer counting (Sr/Ca). We note, however, that R02i constrained by local 'isochrons' assuming binary-mixing relationships should be regarded as helpful local, first-order approximations to delimit the plausible range of initial Th activity ratios, rather than as exact descriptions of the underlying Th sources." (L574ff)

- The authors relied on multiple verifications based on the isochronous method to correct the initial Th. It is known that young, low-thorium samples are highly sensitive to scattering caused by blank errors in the $^{230}Th/^{232}Th$ ratios. Still, the manuscript lacks a good discussion of the contribution of procedural blanks and the propagation of uncertainty. How do the authors handle the blank correction propagating into the uncertainty of the final isochronous line?

We thank Referee #2 for highlighting the importance of blank corrections, especially for young, low-Th samples. We agree that such samples are particularly sensitive to scattering in measured $^{230}Th/^{232}Th$ ratios in general, but possibly also arising from procedural blanks. In the revised manuscript we now report the levels of our procedural blanks for Th and U and add a statement that all ages and activity ratios presented in the paper are corrected for these blanks:

"Chemical blanks are regularly monitored and commonly below 0.4 fg for 234U and 0.04 fg for 230Th (Kerber et al., 2023; Kerber et al., 2025)." (L225)

Notably, the individual procedural blanks analyzed along with the Áaktun Kóopo' series agree with these long-term observations (Kerber et al., 2023; Kerber et al., 2025), with values of 0.04 ± 0.03 (MAD) fg for ^{230}Th, 0.002 ± 0.001 ng for ^{232}Th, and 0.4 ± 0.2 (MAD) fg for ^{234}U. The Áaktun Kóopo speleothems generally have moderate to high U, so the signal-to-noise ratio for most Th measurements is favorable and blank contributions are small compared to the total Th signal. Only for the very youngest samples does the signal-to-noise ratio become poorer and the relative influence of the blank correction increase. For such samples, this contribution can be non-negligible and may increase the scatter in $^{230}Th/^{232}Th$ ratios.

However, for most samples, the influence of procedural blank variability is negligible for the inferred R02i spread. For example, samples E0-C-69 (#10975) and E0-C-73 (#11237) have relatively high S/N ratios but require R02i correction ratios of 11.5 and 41.5 respectively. By artificially doubling the original procedural blank correction, the required R02i to get the same corrected ages would be in the range of 10 and 37, respectively. In summary, while we cannot rule out that undetected outliers in individual chemistry blanks have contributed to some of the data points, and increase the variability in the data set, several lines of evidence suggest that the main patterns in our isochrons and in the inferred, large range of R02i are robust.

- Since not all speleothem samples have clear annual layers, how are the initial $^{230}Th/^{232}Th$ ratios determined for samples lacking chronological striations?

We thank Referee #2 for this question. In our study, annual Sr/Ca layer counting is only one of several possible constraints on initial $^{230}Th/^{232}Th$ ratios and is explicitly used only where clear, laterally continuous cycles are present (i.e. for selected intervals of E1, E8, and E23-3). For samples or stalagmite sections lacking such chronological laminae, R02i is instead constrained using a combination of (i) local isochrons based on groups of closely spaced subsamples, (ii) stratigraphic reasoning to enforce monotonically increasing ages with depth, and (iii) in a few cases, information from modern carbonate and drip-water activity ratios.

In other words, layer counting is not a prerequisite for determining initial $^{230}Th/^{232}Th$; it provides an independent check and local constraint where available. In sections without annual layers we rely on isochrons and stratigraphic constraints alone, and the associated uncertainty in R02i (and thus in the corrected ages) is explicitly propagated into the final age–depth models.

- The manuscript applies two-endmember mixing regressions while simultaneously arguing for highly variable and unsystematic initial thorium compositions. How do the authors reconcile the use of a binary mixing framework with evidence suggesting multiple temporally and spatially variable thorium sources?

We thank Referee #2 for this thoughtful question. We fully agree that, at the cave scale and over longer time spans, Áaktun Kóopo'. speleothems likely receive Th from multiple sources whose contributions vary in space and time. In this context, fitting two-endmember mixing regressions should be understood as a local, first-order approximation applied to subsets of samples that behave approximately linearly in $^{230}Th/^{238}U$ - $^{232}Th/^{238}U$ space, rather than as a literal statement that the entire cave system is governed by only two fixed Th reservoirs.

In practice, the binary-mixing fits are used to obtain approximate R02i values for restricted depth intervals where the data define a reasonably coherent trend, and these estimates are then considered alongside other constraints. In several key intervals, the resulting age–depth models are consistent with independent proxy comparisons, indicating that the two-endmember framework, although undoubtedly simplified, is sufficient as a first-order tool to recover robust chronologies. Our intention is not to deny the existence of multiple Th sources, but to show that, when applied locally and interpreted cautiously, binary mixing still provides useful constraints on initial Th under such heterogeneous conditions.

- The manuscript addresses the limitations of fixed initial thorium corrections under highly "detrital-contaminated" conditions. However, the exclusion of samples with $^{230}Th/^{232}Th$ ratios < 200 raises an important methodological question regarding the practical applicability of the proposed approach. If the most strongly contaminated samples still need to be discarded to achieve robust chronologies, to what extent does the proposed framework truly overcome the limitations associated with high detrital thorium contamination? The current workflow appears to improve uncertainty treatment primarily for moderately contaminated samples rather than resolving the challenges posed by the most problematic end-members.

We thank Referee #2 for this comment and agree that our current framework does not fully resolve the challenges posed by the most strongly detrital-contaminated samples. In this study, we deliberately apply first-order approximations, which work best at improving the treatment of initial Th uncertainty for moderately contaminated samples. Such approaches may help to keep analytical effort and laboratory costs at a feasible level especially for future studies in similar environments. We see our approach rather as a pragmatic compromise than a final solution, and we expect that future studies building on this framework may eventually help to better understand the underlying processes and thus allow more rigorous correction – and thus wider use – of strongly contaminated material as well.

- The manuscript primarily focuses on the propagation of age uncertainties associated with initial thorium contamination. However, potential variability in uranium concentrations may also significantly affect the reliability of U–Th ages, particularly under prolonged flooding or submergence. Although petrographic observations suggest no obvious recrystallization or diagenetic alteration, preservation of primary fabrics alone may not fully exclude subtle uranium remobilization or cryptic open-system behavior. Given the hydrologically dynamic cave environment and evidence for flooding events, how do the authors demonstrate that uranium concentrations and isotopic compositions were not modified post-depositionally?

We thank Referee #2 for this important comment. In principle, prolonged flooding or submergence could promote subtle U remobilization and open-system behavior, but in our case we consider this scenario unlikely for most of the dated material. The majority of stalagmites analyzed in this study come from cave sections that are not regularly flooded and show no evidence for long-lasting inundation. Also petrographic inspection revealed well-preserved primary fabrics without recrystallization or secondary calcite cements. Along individual growth axes, U concentrations and $\delta^{234}U$ values vary largely consistent, which argues against major post-depositional modification of uranium contents or isotopic compositions. However, we acknowledge that this does of course not exclude the possibility of cryptic open-system effects on the more localized scale. However, we interpret the remaining, age inversions as being likely related to a number of possible effects. In addition to possible post-depositional modification on the microscale, undetected micro-hiatuses or uncertainties in initial Th correction (e.g., related to a more complex (isotopic) composition of the end-member(s)) could be possible reasons.

We will add a more explicit statement in section 4.1 of the discussion:

"*Petrographic observations suggest no obvious recrystallization or diagenetic alteration; the stalagmites show well preserved primary fabrics and display largely coherent U concentrations and δ^234U values along their growth axes. Therefore, we consider wide-spread post depositional U remobilization to play a negligible role for our dated material.*" (L326ff)

- Human activity increases the uncertainty of cave research. Did the Maya people's modification of the cave walls affect the cave's sedimentary environment? How to identify non-climate impacts of human activity changes?

We thank Referee #2 for this thoughtful question. Archaeological work in Áaktun Kóopo' has clearly documented that the Maya modified parts of the cave (e.g. low walls and partitions), but these studies have focused on the spatial organization and ritual use of the cave rather than on a detailed reconstruction of its sedimentological evolution. At present there is therefore no quantitative evidence that the Maya constructions systematically altered the large-scale sedimentary environment or hydrology of the chambers from which our stalagmites were collected.

Most stalagmites that contain distinct flooding layers come from a chamber that lies higher than areas that are more regularly flooded (e.g., the mortuary chamber), so inundation of these parts of the cave is expected to reflect comparatively rare, high-water events rather than day-to-day human activity at the cave floor. Nonetheless, human influences on the cave and the overlying soil–vegetation system cannot be ruled out, and non-climatic impacts are difficult to isolate with the available data.

- The manuscript emphasizes that dry conditions have been a major driver of social and environmental change among the Maya. However, much of the subsequent discussion focuses on flood events, inundation, and enhanced clay transport within the cave system. This raises an important question: the relationship between long-term hydroclimatic trends and short-term extreme hydrological events. Can the authors clarify whether the identified flood events are transient, superimposed on an overall drought trend, or reflect a trend of increasing hydroclimatic variability? Which factors contribute to the profound social and political upheaval marked by widespread population decline and the abandonment of many urban centers? A clearer distinction between mean drought conditions and sudden flood dynamics will help strengthen paleoclimate interpretation and improve the coherence of climate narratives.

We thank Referee #2 for this comment and agree that the relationship between long-term hydroclimatic trends, short-term extremes, and Maya societal change is an important topic that has been discussed in literature since a long time. We would like to clarify, however, that the present manuscript does not suggest any causal link between drought conditions and social or environmental change. Where such links are mentioned (e.g., L97ff in the introduction) we explicitly refer to conclusions drawn in previous studies rather than to new evidence from Áaktun Kóopo'. We are convinced that the motivation for this study is phrased sufficiently clear with:

"This study focuses on constructing precise ^{230}Th/U chronologies for stalagmites from Áaktun Kóopo' Cave, […] These chronologies will allow to fill a notable gap in high-resolution paleoclimate records for the region and provide a foundation for future investigations into hydroclimatic and environmental changes, particularly during periods of profound Maya societal changes." (L102ff)

In principle, combining the Áaktun Kóopo' flood layers with the stable isotope and trace-element records could help to explore how short-term extreme events relate to longer-term hydroclimatic variability. However, a rigorous treatment of this question would require a dedicated analysis of proxy sensitivities (e.g. the drivers of $\delta^{18}O$) and a more extensive discussion than is possible within the scope of this methods-oriented study. We therefore refrain from drawing firm conclusions here about whether the identified flood events are transient features superimposed on an overall trend or reflect changes in hydroclimatic variability itself. Instead, we now explicitly note in the Conclusions that the chronologies and proxy records established in this work provide a framework for future studies that will address the relationship between long-term hydroclimate trends, short-term extremes, and their potential relevance for Maya cultural history in more detail.

The last section of the conclusions will be changed to:

"*Ultimately, the improved chronologies and accompanying coherent proxy records from Áaktun Kóopo' Cave now provide a robust foundational framework for future high-resolution reconstructions of hydroclimate and environmental change throughout the timeline of Maya cultural evolution on the YP, e.g., how long-term trends and short-term extreme events, such as the identified flood layers, are related and how they may have intersected with phases of cultural change."*

Response letter – Community comment (Giacomo Medici)

Comments of reviewers in black, author responses in blue

General comments

Very good research on paleo-research with angle on the climate of the past. Please, follow my straightforward comments to improve the manuscript.

Dear Giacomo Medici, we thank for the positive feedback to our manuscript.

Specific comments

Line 14. “Natural formations”. I would use another term like “natural features”. “Formation” is a term used in lithostratigraphy with another meaning.

We will change the wording of the abstract, and the term “formations” is no longer used.

Lines 33-35. “Reconstructing past environmental and climatic variability relies on archives that combine continuity, sensitivity to climate, and robust dating methods. Speleothems fulfil these criteria, as their layered carbonate growth can be precisely dated using U-series disequilibrium dating”. Insert two recent papers on speleothems in carbonate deposits with a paleoclimate link:

- Koltai, G., Scholger, R., Knipping, M., Wang, J., Cheng, H., Frisia, S., ... Spötl, C. 2026. Ancient speleothem giant preserved in a high-Alpine cave (Dolomites, N Italy): Rare insights into the Neogene. Palaeogeography, Palaeoclimatology, Palaeoecology, 113619.

- Medici, G., Marianelli, D., Cornacchia, I., Gori, F., Brandano, M. (2026). Multi-disciplinary approach to paleokarst occurrence in the Eocene–Oligocene succession of the Apulia Carbonate Platform (Salento, Italy). Facies, 72(2), 17.

We thank for the suggestions and will add these were appropriate.

Line 103. Clarify and describe the specific objectives of your research by using numbers (e.g., i, ii, and iii).

We will add numbers to this paragraph.

Line 113. Insert more detail on the depositional paleoenvironment for all the sedimentary deposits described in this research.

In the revised manuscript, we have expanded the description of the cave and sedimentary context in the Study site section (2.1), including the nature of the *kankab* red clay infill, the range from dry to muddy conditions, and the position of the sampled stalagmites relative to water bodies and flooded passages. The speleothem material is described in detail in section 2.2. Since the primary focus of our work is on U–Th systematics and age-model construction rather than on a full sedimentological reconstruction of Áaktun Kóopo’ Cave, we consider this level of detail sufficient to contextualize the samples without diverting from the main scope of the study.

Line 205. “Subsamples”. This terminology is unclear. I suggest a change.

We will change to “*samples*”.

Figures and tables

You might want to add images of the study site, and the specific cave.

At present, we have opted not to include additional photographs of the cave, because we do not have a consistent set of images that would be directly informative for the U–Th dating strategy or the age-model results, and arbitrary views of individual chambers could be misleading rather than helpful. The existing location map and cave plan already show the spatial setting of the sampled stalagmites and drip sites, which we consider sufficient to support the methodological focus of this manuscript.

Figure 1. Issues of graphic resolution. The coordinates are un-readable.
Figure 1. The current version of this figure is not very informative. The scope of Figure 2b is unclear.

We thank for the valuable suggestions to improve Fig.1. We will revise this Figure, e.g. by removing panel b, and improving readability of the coordinates.

Figure 2a. I would add a correlation/regression equation and relative line/curve.

In Figure 2a, our goal is to illustrate the overall first-order mixing trend and the large spread in activity ratios, rather than to derive or use a specific regression for quantitative interpretation. Adding a fitted line and equation would visually dominate the panel without changing any of the subsequent analysis (which is based on local isochrons and stratigraphic constraints), and we therefore prefer to keep the figure focused on the data distribution itself.

Figure 4. I would increase the size of numbers and letters. They are difficult to read.

We appreciate this suggestion. We will increase the font size for the small labels to increase readability.

Figure 4. You might consider inserting a spatial scale for the speleothems.

To clarify the spatial scale of the speleothems relative to the depth axis, we will explicitly state in the Figure 4 caption that the speleothem scans are at the same vertical scale as the y-axis depth (dft) values.

References

Frank, N., Mangini, A., and Korfmann, M.: 230Th/U dating of the Trojan ‘water quarries’, Archaeometry, 44, 305-314, 2002.

Kaufman, A.: An evaluation of several methods for determining 230ThU ages in impure carbonates, Geochimica et Cosmochimica Acta, 57, 2303-2317, 1993.

Kennett, D. J., Masson, M., Lope, C. P., Serafin, S., George, R. J., Spencer, T. C., Hoggarth, J. A., Culleton, B. J., Harper, T. K., Prufer, K. M., Milbrath, S., Russell, B. W., González, E. U., McCool, W. C., Aquino, V. V., Paris, E. H., Curtis, J. H., Marwan, N., Zhang, M., Asmerom, Y., Polyak, V. J., Carolin, S. A., James, D. H., Mason, A. J., Henderson, G. M., Brenner, M., Baldini, J. U. L., Breitenbach, S. F. M., and Hodell, D. A.: Drought-Induced Civil Conflict Among the Ancient Maya, Nature Communications, 13, 3911, 10.1038/s41467-022-31522-x, 2022.

Ludwig, K. R. and Titterington, D. M.: Calculation of (230)Th/U Isochrons, Ages, and Errors, Geochimica Et Cosmochimica Acta, 58, 5031-5042, Doi 10.1016/0016-7037(94)90229-1, 1994.

Moseley, G. E., Richards, D. A., Smart, P. L., Standish, C. D., Hoffmann, D. L., ten Hove, H., and Vinn, O.: Early–middle Holocene relative sea-level oscillation events recorded in a submerged speleothem from the Yucatán Peninsula, Mexico, The Holocene, 25, 1511-1521, 2015.

Richards, D. A. and Dorale, J. A.: Uranium-series chronology and environmental applications of speleothems, Uranium-Series Geochemistry, 52, 407-460, Doi 10.2113/0520407, 2003.

Ridley, H. E., Asmerom, Y., Baldini, J. U. L., Breitenbach, S. F. M., Aquino, V. V., Prufer, K. M., Culleton, B. J., Polyak, V., Lechleitner, F. A., Kennett, D. J., Zhang, M. H., Marwan, N., Macpherson, C. G., Baldini, L. M., Xiao, T. Y., Peterkin, J. L., Awe, J., and Haug, G. H.: Aerosol forcing of the position of the intertropical convergence zone since AD 1550, Nature Geoscience, 8, 195-200, 10.1038/Ngeo2353, 2015.

Stinnesbeck, W., Rennie, S. R., Avilés Olguín, J., Stinnesbeck, S. R., Gonzalez, S., Frank, N., Warken, S., Schorndorf, N., Krengel, T., and Velázquez Morlet, A.: New evidence for an early settlement of the Yucatán Peninsula, Mexico: The Chan Hol 3 woman and her meaning for the Peopling of the Americas, Plos one, 15, e0227984, 2020.