Web links to the author's journal account have been redacted from the decision letters as indicated to maintain confidentiality.

1st Mar 21

Dear Dr Mucek,

Thank you for your patience while your manuscript has been in review. Your manuscript titled "Coupled 40Ar/39Ar and (U-Th)/He thermochronometry reveals resurgence initiation and cold-stored regions of long-lived warm magma at Toba Caldera, Sumatra" has now been seen by 3 reviewers, whose comments are appended below. You will see that they find your work of potential interest. However, the reviewers have raised quite substantial concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication in its present form, but would be interested in considering a revised version that fully addresses these concerns.

We hope you will find the reviewers' comments useful as you decide how to proceed. Should additional work allow you to address these criticisms, we would be happy to look at a revised manuscript. However, please bear in mind that we will be reluctant to approach the reviewers again in the absence of substantial revisions.

Please ensure that a revised version of the manuscript addresses the following editorial concerns:

- detailed explanation and justification of all aspects of your modelling and your statistical approach.

- provide a clear and thorough presentation of your geochronological data set, as requested by Reviewer #3, including a clearer illustration of the data (both Ar-Ar and U-Th/He)and uncertainties in Fig. 1 and inclusion of the data itself as supplementary files.

- ensure that your conclusions and interpretations are fully supported by the data you present, and caveat or tone-down your interpretations where necessary.

- significant improvement in the writing style and tone following suggestions for replacing uncommon words with more widely used synonyms, deleting redundant words and tightening the text.

If implementing these changes in the revision process takes significantly longer than three months, we will be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

We understand that due to the current global situation, the time required for revision may be longer than usual. We would appreciate it if you could keep us informed about an estimated timescale for resubmission, to facilitate our planning. Of course, if you are unable to estimate, we are happy to accommodate necessary extensions nevertheless.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the referees' comments (which should be in a separate document to any cover letter) and any completed checklist:

[link redacted]

** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further. Thank you for the opportunity to review your work.

Best regards,

Derya Gürer, PhD
Communications Earth & Environment
orcid.org/0000-0001-5884-9160

EDITORIAL POLICIES AND FORMAT

If you decide to resubmit your paper, please ensure that your manuscript complies with our editorial policies and complete and upload the checklist below as a Related Manuscript file type with the revised article:

Editorial Policy Policy requirements

For your information, you can find some guidance regarding format requirements summarized on the following checklist:(https://www.nature.com/documents/commsj-phys-style-formatting-checklist-article.pdf) and formatting guide (https://www.nature.com/documents/commsj-phys-style-formatting-guide-accept.pdf).

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

The submission by Adonara Mucek uses coupled thermochronometry to understand the timing of late stage eruptive activity at Toba following the climactic eruption of the Younger Toba Tuff. Overall the paper is well presented, convincing in its arguments, and to a high standard. I cannot see any flaws in the data presentation, nor the conclusions based on these results. The findings are of interest to the geoscience community, so I would recommend publication of this manuscript after minor revision.

The only significant point I would like clarifying is where in the system the authors think that the ZHe-Ar/Ar disequilibrium occurred. Is this thermal window still in the shallow crust, or is it once the lava domes have been subaerially exposed? If it is the latter, given the strong temperature dependence of helium in zircon, would the authors expect to find significant differences in calculated ZHe ages from the edges to the middle of the lava domes? Could future sampling trips do transects of such lava domes to test this hypothesis? Where were the current samples taken from with respect to the lava dome boundaries? I am fairly certain this is an easy point to address.

A smaller point, the writing uses language that is uncommon in scientific papers. I personally like the anecdotes such as the blind men and the elephant, but other examples such as “the afterparty after the big dance” do not add anything to the paper and should be removed. In other instances, very obscure words are used where more widely used synonyms would be better (see below).

Specific Points:
Lines 28-31. This is quite a clunky first sentence. Suggest splitting in two and rewording.
Line 69. Referred to by whom? Suggest removal
Line 75. Flip to “36 to 42 kyrs”
Line 94. The use of “unattended” is not great here. Suggest a different synonym.
Line 129. Same number of significant figures please (i.e. 8.0 ±2.1 kyr)
Line 147. The word “connote” is extremely rarely used, and is likely to make many readers reach for a thesaurus. Suggest replacing with “implies”, “indicates”, or “suggests”
Line 151. Define what temperature range one would expect for a “cold halo”. I know the article goes on to discuss these temperature bounds, but it is worth adding here too.
Line 155. Write the full element name when starting a sentence (i.e. Helium)
Line 188: Replace “posit” with “propose”
Line 231: “climactic” instead of “climatic”
Line 275: Should be “eruptions” plural for the domes

Reviewer #2 (Remarks to the Author):

In this paper, the authors build on an earlier paper by mostly the same authors to further ascertain the temporal (and thermal) evolution of post-YTT lava domes. The topic of renewed eruptive activity in a caldera setting is an important one to address, the dataset is state-of-the-art and sound, and the paper is quite well written and organized. The thermal modelling approach is valuable but not well enough explained for non-initiates and therefore does not convey as much information as it could. The outcomes are interesting but not particularly new in that they mainly corroborate and strengthen the earlier paper by Mucek et al. (2017) in Nature Communications. Please see below specific comments on the text, including comments to help improve the section on thermal modelling.

Line 17-19: Connote is used twice and feels a bit repetitive
Line 21: “Assuming that these domes are at the origin of the resurgent uplift of Samosir island, this would have initiated…”
Line 68-69: Not sure the “Colloquially referred to as the ‘afterparty after the big dance’” is needed.

Line 72: It’s not the magma that rebounds. Maybe topography or structure would be better?
Line 76: Again, “rebounding” is not an ideal term. The influx of magma from depth might be continuous from pre- to post-caldera formation (it might even be less post- than pre-caldera formation according to some studies). Continuous influx from depth gradually re-inflates the reservoir that had been emptied during the caldera-forming eruption, hence the crust elevation ‘rebounds’.
Line 84: “Other post-YTT [vents/eruptives] include…”
Line 145: A brief explanation of the philosophy behind the thermal modelling would be useful. What exactly does the HeFTy program compute? Given the information provided in supplementary material, it considers the diffusive loss (or retention) of Ar and He in feldspar and zircon, respectively, but how does it compute the cooling trajectories? There is some explanation on the input parameters (although it needs to be reorganized a bit) and no information about what exactly is computed. It’s useful to have a short explanation in the main body of the text too, and not only in the supplementary material.
Line 205: re-set the Ar clock in sanidine (the ZHe is re-set).
Line 211: Delete “among many others”
Line 22: “represented by the Samosir domes”
Line 261-263: Not the best case, or example. It would be erroneous to interpret pre-YTT magma storage conditions from post-YTT domes. It would be comparing apples and oranges (Or adult and baby elephants maybe).
Line 422: “found in Supplementary”

Supplementary File:
Line 32: planchettes
Line 156: “such as, for example” – unnecessary repetition
Line 251: A brief description is needed, what is the input data? What does the model compute?
Line 264: Some constraints must have been chosen to run the model, what were they?
Line 269: “in the domes the zircon”
Line 308-323: This paragraph should come at the beginning of Supplementary Information 5, before the presentation of the 4 modelling approaches, to provide context for them. How is the input data used?
Line 318: “to support another assumption”
Line 348-349: A-B-C-D only, no E
Line 370: swapped, (ii) is ZHe.
Line 398-399: What is computed to constrain this?
Line 412-413: No light green lines visible in the figures

Reviewer #3 (Remarks to the Author):

This paper reports new feldspar 40Ar/39Ar and zircon U-Th/He dates from 5 samples including two from the enormous younger Toba Tuff, and three from minute post-Toba Tuff domes. The domes are characterized by U-Th/He dates that appear to be slightly younger than the 40Ar/39Ar dates from all 5 samples. The authors suggest that this reflects post-caldera emplacement of relatively cool rhyolite from the roof zone of the Toba magma reservoir that permitted retention of argon accumulated since ca. 74 ka, but that cooled through the closure temperature for retention of He in zircons several thousand years after 74 ka. In my opinion, the data set is rather small to be making such an inference, but that this is a ptentially novel use of this pair of geochronometers to parse out details of magmatic conditions and processes associated with large rhyolitic magma systems. I conclude that the main tenent is worthy of publishing here, but that the paper could benefit from a substantial overhaul.

There are three issues that I would want to see addressed in an improved manuscript:

1. Since the premise rests entirely on the quality of the geochronogic data, show me the data! The schematic probability density curves and histograms in figures 1 and 2 are likely hiding a great deal of complexity of the sets of dates acquired from whithin each of the 5 samples. The text and tables in the supplementary documents do not address this complexity. How many crystals were measured? Were there any outliers? Was any filtering of dates necessary? Whether or not filtering was necessary, what are the implications of any skewness? The individual dates, with their associated uncertainties, should be portrayed in figure 1 in the main body of the paper, and somehow portrayed against the statistically derived histograms in Figure 2. This extends to details of the published 40Ar/39Ar data for the YTT that are used uncritically . I am not so sure I would weight the poorer quality data in the Mark et al. paper on equal footing with the data in Storey et al. If one used only the more precise, and potentially more accurate, dates in Storey et al., would the delta t values between the helium dates from the domes and their Ar date remain as large as you show?

Further, it is not clear from reading the paper (including the supplementary materials) why a Bayesian statistical analysis is necessary to reveal distinctions in age among the 10 sets of dates from the 5 samples. In part this reflects my point above that the actual data are not laid out for all to see and wander through graphically. The only posterior I can think of is that you know the domes to be younger than the Tuff -- or did I miss something? In my opinion, a much clearer case must be made that the Bayesian approach is essential to the story as it is promoted as one of the novel elements of this work. How would/could your data be (mis-) interpreted if the Bayesian modeling was not undertaken? Are there other example studies that could be cited in which the Bayesian approach was instrumental in reaching an otherwise unattainable conclusion?

2. There is a wider literature and context for the cool vs. warm debate for large silicic magmatic systems. The authors are not the first to conclude that this debate is somewhat ridiculous, and that it is expected to find both cool and warm domains in large systems that coexist, for example:

Andersen, N. L., Singer, B. S., & Coble, M. A. (2019). Repeated rhyolite eruption from heterogeneous hot zones embedded within a cool, shallow magma reservoir. Journal of Geophysical Research: Solid Earth, 124(3), 2582-2600.

Other well documented examples from Taupo and elsewhere could be cited. My point is that a broader context for the finding here that there was a significant thermal gradient in the Toba magma body during and and for several kyr following the caldera-forming eruption, would strenghten the impact of your work and help it influence thought on this issue accross the volcanological community.

3. There are published experimental data on diffusion parameters for argon in sanidine, indicating closure temperatures of ca. 325 oC or greater are possible:

Cassata, W. S., & Renne, P. R. (2013). Systematic variations of argon diffusion in feldspars and implications for thermochronometry. Geochimica et Cosmochimica Acta, 112, 251-287.

I suggest that a the paper would garner more impact if it also included a sensitivity test of the cooling histories shown in Figure 3 that factors in the variation in the kinetic parameters in the above paper, as well as the assumed rate of cooling and crystal (diffusion domain) sizes assumed in the thermal modeling.

More minor concerns are:

1. the introduction is a bit shrill for my taste, and could benefit from taking the tone down a bit. For example, many earth scientists would argue that asteroid impacts and flood basalts like the Siberian or Deccan Traps are equally, if not far more so, catastrophic than rhyolitic super-eruptions. In my opinion the paper is relatively clear and easy to read, but could benefit by another level of strong editing to avoid redundant use of some words (connotes, caldera, etc.) and strive for a more concise read. For example in lines 102-105, the text could be shortened to: "These are concordant with Ar ages of X and Y on sanidine from the YTT, as well as our new Ar ages reported in Supplementary Data item 2".

2. The text below the two panels in Figure 4 is too much. Why can't most of this be placed into the caption and use only key bullets or better yet careful annotations in the cartoons themselves?

Brad Singer
Madison, Wisconsin
February 4, 2021

Response to Reviewers

(**Our responses in blue**)

Reviewer #1 (Remarks to the Author):

The submission by Adonara Mucek uses coupled thermochronometry to understand the timing of late stage eruptive activity at Toba following the climactic eruption of the Younger Toba Tuff. Overall the paper is well presented, convincing in its arguments, and to a high standard. I cannot see any flaws in the data presentation, nor the conclusions based on these results. The findings are of interest to the geoscience community, so I would recommend publication of this manuscript after minor revision.

The only significant point I would like clarifying is where in the system the authors think that the ZHe-Ar/Ar disequilibrium occurred. Is this thermal window still in the shallow crust, or is it once the lava domes have been subaerially exposed?

We believe the thermal window for ZHe-Ar/Ar disequilibrium occurred in the subsurface prior to eruption. The key point here is that the closure temperatures are subsolidus but the transition between the Ar closure temperature to He closure temperature happens over 5 to 15 ka. Thus a thermal window of ~200°C has to be traversed over several thousand years. This time period is orders of magnitude longer than the cooling history on eruption - our samples come from the surface carapaces of the domes so they would have cooled to ambient within a few tens of years based on historic eruptions. We would suggest that the disequilibrium happens in the "cold" (read subsolidus) halo of the magma "chamber".

If it is the latter, given the strong temperature dependence of helium in zircon, would the authors expect to find significant differences in calculated ZHe ages from the edges to the middle of the lava domes?

As we explain above, the thermal window is prior to eruption. On the point here, given that (i) the uncertainty of the ZDD ages is rather large, (ii) the domes are of relatively small size (<2 km in diameter), (iii) the domes are not significantly eroded, hence they were sampled on their peripheral parts and (iv) cooled very quickly, we would be surprised to see an age gradient from the edges to the middle part of the domes. Maybe some drill core data might be a way to test this. Available studies indicate that the out 10's of meters of such domes cool to ambient in a few 10 of years – steam emanation stops with a few years.

Could future sampling trips do transects of such lava domes to test this hypothesis?

Following on from above this is an interesting suggestion that might be tested with drill core data where the deeper parts and conduits have been intersected. There are few examples (Unzen, Coso, and Mono craters) where drill has been done, but they are all too young to use the thermochronology. It might be an interesting project to drill into older domes like those at Toba to test the idea. But they cannot be too old or the resolution of the dating techniques would be exceeded. Toba is ideal in this sense - serendipity.

Where were the current samples taken from with respect to the lava dome boundaries?

Given the limited outcropping conditions, the samples were collected from the external carapaces of the domes. If not from the outermost surface no more than a few meters from the surface of the domes. There is simply no exposure into the interior of the domes yes unless it is man-made (quarries) and then no more than a few tens of meters at most.

I am fairly certain this is an easy point to address.

A smaller point, the writing uses language that is uncommon in scientific papers. I personally like the anecdotes such as the blind men and the elephant, but other examples such as "the afterparty after the big dance" do not add anything to the paper and should be removed.

Apologies, sometimes it is easy to get carried away trying to make the prose more entertaining - removed and rephrased

In other instances, very obscure words are used where more widely used synonyms would be better (see below).

Specific Points:

Lines 28-31. This is quite a clunky first sentence. Suggest splitting in two and rewording. - we have re-stated these.

Line 69. Referred to by whom? Suggest removal – removed, rephrased

Line 75. Flip to "36 to 42 kyrs" - corrected

Line 94. The use of "unattended" is not great here. Suggest a different synonym. – changed to "has not been addressed"

Line 129. Same number of significant figures please (i.e. 8.0 ±2.1 kyr) - corrected

Line 147. The word "connote" is extremely rarely used, and is likely to make many readers reach for a thesaurus. Suggest replacing with "implies", "indicates", or "suggests" – changed to "implies" – we subscribe to the tenet that people "suggest" and data "indicate" that is why use "connote". But we changed to "implies" anyway.

Line 151. Define what temperature range one would expect for a "cold halo". I know the article goes on to discuss these temperature bounds, but it is worth adding here too.

We think it is too early to "give away" the temperature range at this point because we model (and constrain) the temperature range in the next section. The thermal aureole to intrusions ranges from the solidus to hydrothermal temperatures based on metamorphic studies. So 700 to 150°C – the range we are interested in is 500° to 150°C.

Line 155. Write the full element name when starting a sentence (i.e. Helium) - corrected

Line 188: Replace "posit" with "propose" – corrected, although posit is correct.

Line 231: "climactic" instead of "climatic" – corrected.

Line 275: Should be "eruptions" plural for the domes - corrected

Reviewer #2 (Remarks to the Author):

In this paper, the authors build on an earlier paper by mostly the same authors to further ascertain the temporal (and thermal) evolution of post-YTT lava domes. The topic of renewed eruptive activity in a caldera setting is an important one to address, the dataset is state-of-the-art and sound, and the paper is quite well written and organized. The thermal modelling approach is valuable but not well enough explained for non-initiates and therefore does not convey as much information as it could.

We have tried to correct this and expanded our explanations. We added an additional section covering the philosophy behind the modelling and the explanation of the HeFTy modelling principles to the thermal history modelling section as we feel that this part of the manuscript (Interpretation and Discussion) is not appropriate to explain these basics. Additional information was provided in the Supplementary materials as well.

The outcomes are interesting but not particularly new in that they mainly corroborate and strengthen the earlier paper by Mucek et al. (2017) in Nature Communications.

We would respectfully point to the significant advance from our 2017 work that was geochronological. Here we develop (i) a multi-system dating approach consisting of three independent geo-/thermochronometers complemented with (ii) thermal modelling and (iii) Bayesian inference that allow the pre-eruptive magma storage to be addressed. This is novel and allows a rare temperature-time perspective to be applied to pre-eruptive sub-solidus storage of erupted lava, where previous works primarily provide a thermal perspective.

Please see below specific comments on the text, including comments to help improve the section on thermal modelling.

Line 17-19: Connote is used twice and feels a bit repetitive – corrected, one "connote" changed to "implies"

Line 21: "Assuming that these domes are at the origin of the resurgent uplift of Samosir island, this would have initiated…" – We have rephrased this sentence.

Line 68-69: Not sure the "Colloquially referred to as the 'afterparty after the big dance'" is needed. – deleted – yes, we got carried away trying tom make the prose more entertaining for non-specialists.

Line 72: It's not the magma that rebounds. Maybe topography or structure would be better?

Magma rebound is an established term in the caldera literature since Marsh, 1984. The term was coined to describe one of the processes that motivates resurgent uplift. The term describes the inflow of magma that was initially displaced away by the subsiding caldera block. A new lithostatic, magmastatic, and isostatic equilibrium is re-established by displaced magma flowing back into its original space uplifting the caldera floor in the process. We proposed that this was the process occurring at Toba in our 2015 paper (de Silva et al., 2015 Frontiers).

Line 76: Again, "rebounding" is not an ideal term. The influx of magma from depth might be continuous from pre- to post-caldera formation (it might even be less post- than pre-caldera formation according to some studies). Continuous influx from depth gradually re-inflates the reservoir that had been emptied during the caldera-forming eruption, hence the crust elevation 'rebounds'.

See above. There are certainly cases where recharge (influx of magma depth) dominates the post collapse magma history, but in others like Toba and the largest Andean examples we have worked on, remnant magma dominates the record. Recharge is almost certainly occurring but not at a flux that account for uplift, thus a role for magma rebound in the Marsh sense is clear. We agree that this is a poorly understood area of caldera evolution but there are now several studies that have addressed this.

Line 84: "Other post-YTT [vents/eruptives] include…" – changed to "eruptives"

Line 145: A brief explanation of the philosophy behind the thermal modelling would be useful. What exactly does the HeFTy program compute? Given the information provided in supplementary material, it considers the diffusive loss (or retention) of Ar and He in feldspar and zircon, respectively, but how does it compute the cooling trajectories? There is some explanation on the input parameters (although it needs to be reorganized a bit) and no information about what exactly is computed. It's useful to have a short explanation in the main body of the text too, and not only in the supplementary material.

We added an additional section covering the philosophy behind the modelling and the explanation of the HeFTy modelling principles to the Thermal history modelling section as we feel that this part of the manuscript (Interpretation and Discussion) is not appropriate to explain these basics. Additional information was provided in the Supplementary materials as well.

Line 205: re-set the Ar clock in sanidine (the ZHe is re-set). – corrected, thanks for spotting this!

Line 211: Delete "among many others" - corrected

Line 22: "represented by the Samosir domes" - corrected

Line 261-263: Not the best case, or example. It would be erroneous to interpret pre-YTT magma storage conditions from post-YTT domes. It would be comparing apples and oranges (Or adult and baby elephants maybe).

Here we are making the case that the full context of the eruption needs to be taken into account, then sampling and analytical strategy. Clearly we did not do a good job. We have now modified and expanded the text to try and better express our point.

Line 422: "found in Supplementary" - corrected

Supplementary File:

Line 32: planchettes - corrected

Line 156: "such as, for example" – unnecessary repetition - corrected

Line 251: A brief description is needed, what is the input data? What does the model compute? – brief description added, input data described in general.

Line 264: Some constraints must have been chosen to run the model, what were they?

The statement "no time-temperature constraints were applied between the starting and end points" is correct. We define the starting and end points below (i.e., T= 950°C at 76±1 ka and T= 15°C at 0 Ma, respectively), therefore we would like to keep it.

Line 269: "in the domes the zircon" - corrected

Line 308-323: This paragraph should come at the beginning of Supplementary Information 5, before the presentation of the 4 modelling approaches, to provide context for them. How is the input data used?

We believe that with the addition of a new text explaining the modelling principles and basic input data requirements, the current structure is more readable. Readers could be lost after reading all the details about the input parameters before they actually see the purpose and strategy for the modelling. Therefore our preference is to keep the current structure.

Not sure we understand the last question – the input data are inserted/typed into the desired fields in the HeFTy software. HeFTy Installation package contains an excellent manual with detailed user guidelines.

As we explain above, we added an additional section covering the philosophy behind the modelling and the explanation of the HeFTy modelling principles to the Thermal history modelling section as we feel that this part of the manuscript (Interpretation and Discussion) is not appropriate to explain these basics. Additional information was provided in the Supplementary materials as well. We believe these additions will help the reader.

Line 318: “to support another assumption” - corrected

Line 348-349: A-B-C-D only, no E - corrected

Line 370: swapped, (ii) is ZHe. - corrected

Line 398-399: What is computed to constrain this? – rephrased and an additional explanation added

Line 412-413: No light green lines visible in the figures – corrected, should read “blue”

Reviewer #3 (Remarks to the Author):

This paper reports new feldspar 40Ar/39Ar and zircon U-Th/He dates from 5 samples including two from the enormous younger Toba Tuff, and three from minute post-Toba Tuff domes. The domes are characterized by U-Th/He dates that appear to be slightly younger than the 40Ar/39Ar dates from all 5 samples. The authors suggest that this reflects post-caldera emplacement of relatively cool rhyolite from the roof zone of the Toba magma reservoir that permitted retention of argon accumulated since ca. 74 ka, but that cooled through the closure temperature for retention of He in zircons several thousand years after 74 ka. In my opinion, the data set is rather small to be making such an inference, but that this is a ptentially novel use of this pair of geochronometers to parse out details of magmatic conditions and processes associated with large rhyolitic magma systems. I conclude that the main tenent is worthy of publishing here, but that the paper could benefit from a substantial overhaul.

Thank you for the encouragement.

There are three issues that I would want to see addressed in an improved manuscript:

1. Since the premise rests entirely on the quality of the geochronogic data, show me the data! The schematic probability density curves and histograms in figures 1 and 2 are likely hiding a great deal of complexity of the sets of dates acquired from whithin each of the 5 samples. The text and tables in the supplementary documents do not address this complexity. How many crystals were measured? Were there any outliers? Was any filtering of dates necessary? Whether or not filtering was necessary, what are the implications of any skewness? The individual dates, with their associated uncertainties, should be portrayed in figure 1 in the main body of the paper, and somehow portrayed against the statistically derived histograms in Figure 2.

We clearly did not do as good a job of this as we thought, and we have made a concerted effort to address this. We prepared a new figure with rank order plots of all single grain ZHe data and single Ar analyses that are considered in this study. However, given the size and complexity of the figure (we apologize if our creative and graphic skills are unable to think of a simpler and more effective way of presenting these data) we prefer to present it in the supplementary materials and not as a part of the Figure 1 and 2 as requested by the reviewer, simply because showing so many analytical points would defeat the purpose of these two figures and would make them difficult to read and decipher for the general readership. For the geochronologically-savvy, the supplementary information should now provide all the data they need.

We are confused by the criticism regarding the complexity of the data, outliers and data filtering not being represented in the supplementary data – we were careful to ensure these were all reported in the submitted version of the manuscript (Supplementary Materials), but we made some adjustments to improve the readability of the tables.

This extends to details of the published 40Ar/39Ar data for the YTT that are used uncritically. I am not so sure I would weight the poorer quality data in the Mark et al. paper on equal footing with the data in Storey et al. If one used only the more precise, and potentially more accurate, dates in Storey et al., would the delta t values between the helium dates from the domes and their Ar date remain as large as you show?

We are puzzled by this comment – we do not use either Mark et al. nor Storey et al. data to calculate delta t. Maybe we didn't make it clear that the ^{40}Ar-^{39}Ar ages presented in the paper are our own data that are concordant with the two aforementioned data sets. We have gone through the manuscript and we found several places where we clearly state this.

To be clear what we do is the following: First we show that our Ar ages on YTT and post-YTT domes are indistinguishable from the currently accepted Ar age of YTT eruption based on Mark et al. and Storey et al. datasets (i.e., our ages and published ages overlap within uncertainty). This shows that our Ar data are accurate and precise. We then compare our Ar data with our He data obtained on the same samples and based on the statistical evaluation we conclude that there is a credible difference between the Ar and He ages.

However, we find the reviewers comment to be important. So, in order to test the hypothesis that the age of Storey et al. is the most accurate and precise age for the YTT event (despite being disputed by Mark et al. 2012), we compared this age with our He data. We found that delta t is slightly lower than the delta t values calculated for our Ar data, however, regardless of the smaller delta t, there is a credible difference between our He ages and the Ar age of Storey et al. 2012 (please see the plot showing posterior difference of the means). This result proves that our conclusions are valid – i.e., there is a statistically distinguishable difference between He and Ar ages!

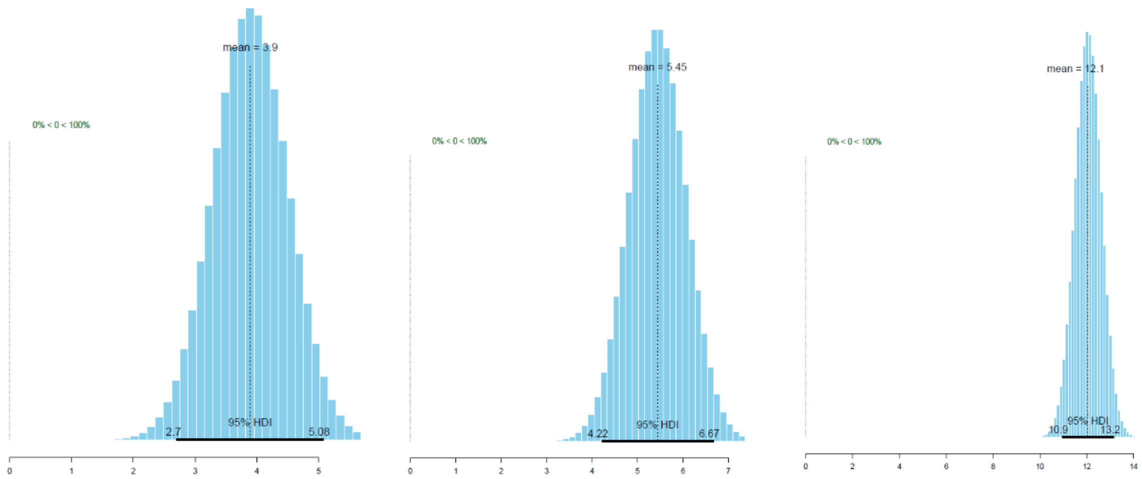


Fig. 1: Posterior distribution of the difference (in kyr) between the mean ZHe ages and the YTT Ar age of Storey et al. (2012). From the left: Samosir North (LT12_003), Samosir (Tuk Tuk) (LT12_008), Paredpur (South) (LT12_015). Note that in all three cases there is a 100% probability of credible difference between the ZHe and the Ar ages, confirming our original interpretation based on the same comparison of our new Ar data and our published ZHe data.

Further, it is not clear from reading the paper (including the supplementary materials) why a Bayesian statistical analysis is necessary to reveal distinctions in age among the 10 sets of dates from the 5 samples. In part this reflects my point above that the actual data are not laid out for all to see and wander through graphically.

A great point which we accept – we added a sentence explaining why the Bayesian inference is so helpful – as recently demonstrated by Morgan et al. (2019, Geology), Bayesian inference allowed to discriminate intervals as short as 20 kyrs in eruptive history of Fish Canyon tuff and associated units (these are ca. 28 Ma old and anaytical uncertainties are in excess of 100 ka), which was not otherwise possible using frequentist approaches. We developed a similar approach which is specifically tailored to compare two populations. The raw data are now presented in Supplementary Figure 3.

The only posterior I can think of is that you know the domes to be younger than the Tuff -- or did I miss something?

With all our due respect, it appears that the reviewer missed something: In Bayesian inference there is no need to consider any posterior (posterior probability is the probability of an event occurring after taking into consideration new information; in other words, it is the outcome of our statistical evaluation). If the reviewer meant "prior" expectation that was used to develop our approach, there are several options. For example, we could have assumed that the domes are younger than YTT eruption because they overlie the YTT tuff and there is sediment between them (and hence must be younger), or we could have assumed that ZHe ages must be younger than Ar ages given their lower closure temperature. However, neither of these plausible prior expectations were applied because we aimed to make an unbiased comparison of two independent groups of data in order to achieve the most "honest", unbiased results. As we write in the paper (L.342 to 344), the only uninformative prior expectation used was that ages cannot be negative (i.e., from the future) and that expected distribution is normal.

In my opinion, a much clearer case must be made that the Bayesian approach is essential to the story as it is promoted as one of the novel elements of this work.

We believe that with the addition of the reference to Morgan et al. (2019) paper and the explanation of the advantages of the Bayesian approach, this case is now stronger. We do present a novel approach to comparison of two datasets written in r-code that can be used by other practitioners, and this is one of the novel elements of this work.

How would/could your data be (mis-) interpreted if the Bayesian modeling was not undertaken?

A classical approach would be a two sample t-test for comparison of two groups, however, this "frequentist" approach is inappropriate to be used from the beginning (hence we do not use it) at least for the following three reasons: (i) it does not consider uncertainties of data-points (which is really critical for geochronological data with overlapping uncertainties), (ii) it is based on p-value which is rather arbitrarily defined and allows only "yes" or "no" type of conclusion, (iii) it assumes the variances for the two independent groups are equal (clearly not the case for geochronological data). In contrast, our Bayesian inference for comparison of two groups takes into account uncertainties of individual data-points, does not assume equal variance of the two groups, and it allows to calculate probability of the difference between two groups.

Are there other example studies that could be cited in which the Bayesian approach was instrumental in reaching an otherwise unattainable conclusion?

- Yes, we now refer to paper by Morgan et al. (2019, Geology) where the authors demonstrated that by using Bayesian inference it is possible to discriminate time intervals of 20 kyrs for eruption ages of ~28 Ma.

2. There is a wider literature and context for the cool vs. warm debate for large silicic magmatic systems. The authors are not the first to conclude that this debate is somewhat ridiculous, and that it is expected to find both cool and warm domains in large systems that coexist, for example:

Andersen, N. L., Singer, B. S., & Coble, M. A. (2019). Repeated rhyolite eruption from heterogeneous hot zones embedded within a cool, shallow magma reservoir. Journal of Geophysical Research: Solid Earth, 124(3), 2582-2600.

We agree and recognise the reviewer's group is one of the few that have challenges this somewhat specious debate successfully. We did already cite this group's 2017 work, but we will also cite this. We note that our approach is different as it is a paired thermochronology approach but are converging on similar conclusions. The more studies that address this "debate" (and we do agree with the reviewer's characterisation - but it is a popular theme), the less of a band-wagon it will become.

Other well documented examples from Taupo and elsewhere could be cited. My point is that a broader context for the finding here that there was a significant thermal gradient in the Toba magma body during and for several kyr following the caldera-forming eruption, would strenghten the impact of your work and help it influence thought on this issue accross the volcanological community.

This is a valuable recommendation . We are familiar with the Taupo literature, but none of the studies we are aware of provide a quantitative thermotemporal approach to the pre-eruptive storage that informs the warm vs cold debate. However, we do find that Barker et al., 2014 is a useful reference for thinking about pre-eruptive evolution of post-super eruption magmatic evolution. We now refer to this and add a short sentence to the manuscript.

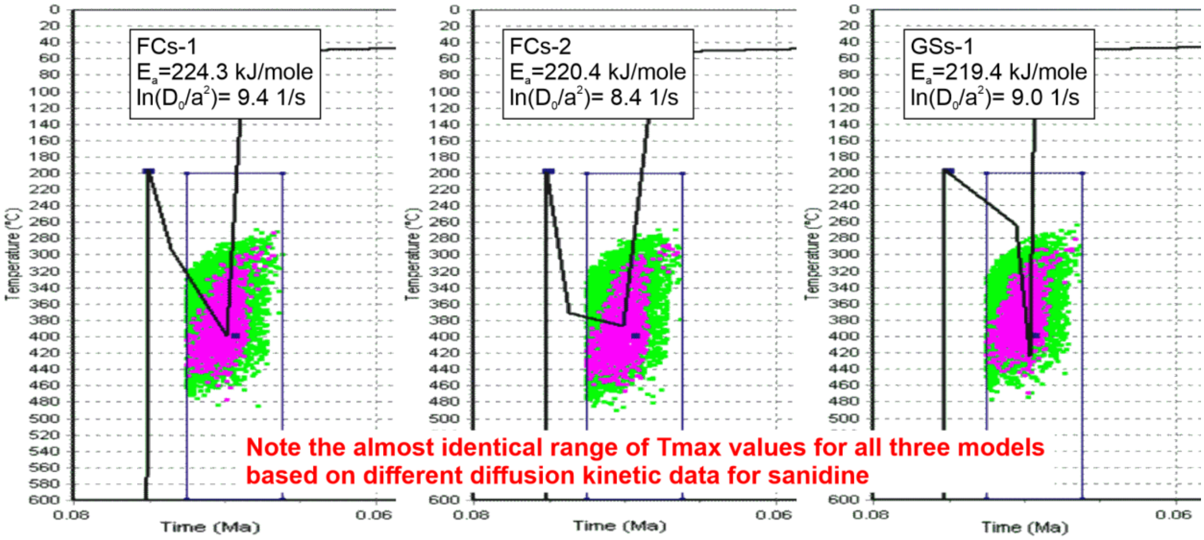
3. There are published experimental data on diffusion parameters for argon in sanidine, indicating closure temperatures of ca. 325 oC or greater are possible:

Cassata, W. S., & Renne, P. R. (2013). Systematic variations of argon diffusion in feldspars and implications for thermochronometry. Geochimica et Cosmochimica Acta, 112, 251-287.

I suggest that a the paper would garner more impact if it also included a sensitivity test of the cooling histories shown in Figure 3 that factors in the variation in the kinetic parameters in the above paper, as well as the assumed rate of cooling and crystal (diffusion domain) sizes assumed in the thermal modeling.

Firstly, as we state in the supplementary materials, for this study we did use diffusion parameters from the recommended study by Cassata and Renne (2013). These authors report three very similar values for sanidine which overlap within 2 sigma uncertainty (closure temperatures are 330±8°C, 332±11°C, 322±6°C), therefore it is unlikely that different diffusion parameters will affect significantly our modelling results. To test this hypothesis, we re-modelled the sample from Samosir North – which is the one that should be affected the most, using the three different diffusion coefficients. As shown below, there is indeed no discernible difference between the models. Therefore, we conclude that different diffusion coefficients reported by Cassata and Renne (2013) have no significant impact on our results.

Thermal modelling results for sample Samosir North based on three different diffusion kinetic data reported for sanidine by Cassata & Renne 2013



Secondly, (i) there is no need to assume a cooling rate in our thermal modelling approach and (ii) there were no constraints imposed on cooling rate in the modelling as we state in the Supplementary file (L.334: “No limit on cooling rate between the t-T constraints was imposed.”). Thirdly, crystal (diffusion domain) sizes assumed for the modelling are based on the measured dimensions of analysed zircon and feldspar crystals (which we consider to represent the size of diffusion domains, which is definitely a valid assumption for (U-Th)/He system (Reiners et al. 2004) and very likely a valid assumption for sanidine Ar system (Cassata and Renne, 2013). We believe that this approach based on real measurements is more appropriate than an “assumption” based approach, therefore we would like to keep it.

Given the three arguments presented above, we do not think that there is a need for a sensitivity test of the modelling approach and therefore we have not included it into the paper.

More minor concerns are:

1. the introduction is a bit shrill for my taste, and could benefit from taking the tone down a bit. For example, many earth scientists would argue that asteroid impacts and flood basalts like the Siberian or Deccan Traps are equally, if not far more so, catastrophic than rhyolitic super-eruptions. – corrected, we added “some of the most catastrophic...”

In my opinion the paper is relatively clear and easy to read, but could benefit by another level of strong editing to avoid redundant use of some words (connotes, caldera, etc.) and strive for a more concise read. For example in lines 102-105, the text could be shortened to: "These are concordant with Ar ages of X and Y on sanidine from the YTT, as well as our new Ar ages reported in Supplementary Data item 2". – corrected accordingly

2. The text below the two panels in Figure 4 is too much. Why can't most of this be placed into the caption and use only key bullets or better yet careful annotations in the cartoons themselves?

Yes, we have taken care of this.

12th Jul 21

Dear Dr Mucek,

Please allow me to apologise for the delay in sending a decision on your revised manuscript titled "Coupled 40Ar/39Ar and (U-Th)/He thermochronometry reveals resurgence initiation and cold-stored regions of long-lived warm magma at Toba Caldera, Sumatra.". It has now been seen by our original reviewers, whose comments appear below. In light of their advice I am delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment under the open access CC BY license (Creative Commons Attribution v4.0 International License).

We therefore invite you to revise your paper one last time to address the remaining concerns of our reviewers. Specifically, please

- move Supplementary Figure 1 into the main body (as suggested by reviewer 3) along with an explanation of how Bayesian statistical methods yield the very simple looking normal distributions shown in Fig. 1B and Fig. 2.

- expand discussion of radiometric data and discuss implications for the claim of a significant age difference between the ZHe dates from 3 domes, and the Ar and ZHe dates from the YTT samples.

- check whether the captions of Supplementary Figures 4 and 5 are swapped, and correct this if necessary.

At the same time we ask that you edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table". Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at https://www.nature.com/documents/commsj-file-checklist.pdf .

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication under a CC BY license (Creative Commons Attribution 4.0 International License). This license allows maximum dissemination and re-use of open access materials and is preferred by many research funding bodies.

For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit https://www.nature.com/commsenv/about/open-access

At acceptance, the corresponding author will be required to complete an Open Access Licence to Publish on behalf of all authors, declare that all required third party permissions have been obtained and provide billing information in order to pay the article-processing charge (APC) via credit card or invoice.

Please note that your paper cannot be sent for typesetting to our production team until we have received these pieces of information; therefore, please ensure that you have this information ready when submitting the final version of your manuscript.

Please use the following link to submit the above items:
[link redacted]
** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Derya Gürer, PhD
Editorial Board member
Communications Earth & Environment
orcid.org/0000-0001-5884-9160

Heike Langenberg, PhD
Chief Editor
Communications Earth & Environment

On Twitter: @CommsEarth

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

The revised manuscript by Adonara Mucek and co-authors is a significant improvement on an already polished first submission. The authors have addressed all of my minor points satisfactorily, and responded well to the more critical reviewer 3 as well. I am satisfied with the changes that have been made and think that this is now ready for publication. I have no further comments and suggestions.

Reviewer #2 (Remarks to the Author):

Review 2: Coupled 40Ar/39Ar and (U-Th)/He thermochronometry reveals resurgence initiation and cold-stored regions of long-lived warm magma at Toba Caldera, Sumatra

It is the second time I review this paper and I found that the authors thoroughly addressed all the reviewers' comments. In this paper, the authors build on an earlier paper by mostly the same authors to further constrain the temporal and thermal evolution of post-YTT lava domes. The topic of renewed eruptive activity in a caldera setting is an important one to address, the dataset is state-of-the-art and sound, and the paper is well written and organized. The Bayesian and thermal modelling approaches are valuable and are now better explained. I thus recommend it for publication. I just noticed that the captions of Supplementary Figures 4 and 5 appear swapped.

Reviewer #3 (Remarks to the Author):

Thank you for preparing Figure 1 of the revised supplementary documents. This figure vividly demonstrates my main concern with the interpreted ZHe and Ar ages from the 5 samples. In my opinion this figure simply must be presented in the main body of the paper, and a more robust explanation of how Bayesian statistical methods are able to cull these data sets down to yield the very simple looking normal distributions shown in Fig. 1B and Fig. 2 of the paper must be provided. The impact of this paper will rest not only on the quality of the new isotopic data, but more importantly on how such a highly complex, and largely overlapping set of individual ZHe and Ar dates from each sample are interpreted.

See for example a recent paper by Schaen, A.J., Jicha, B.R., Hodges, K.V., Vermeesch, P., Stelten, M.E., Mercer, C.M., Phillips, D., Rivera, T.A., Jourdan, F., Matchan, E.L. and Hemming, S.R., 2021. Interpreting and reporting 40Ar/39Ar geochronologic data. GSA Bulletin, 133, pp.461-487.

This would provide a transparent look into support for the claim of a significant age difference between the ZHe dates from 3 domes, and the Ar and ZHe dates from the YTT samples.

I am supportive of publication provided these points are addressed.

RESPONSE TO EDITOR'S AND REVIEWERS COMMENTS

Reviewer 3 comments:

Thank you for preparing Figure 1 of the revised supplementary documents. This figure vividly demonstrates my main concern with the interpreted ZHe and Ar ages from the 5 samples. In my opinion this figure simply must be presented in the main body of the paper,

- agreed and accepted, we added the Figure 1 from the supplementary documents into the main text (now as Figure 2) to show the raw data and offer a more transparent look into the data, which we believe support our case regardless.

and a more robust explanation of how Bayesian statistical methods are able to cull these data sets down to yield the very simple looking normal distributions shown in Fig. 1B and Fig. 2 of the paper must be provided.

- this may be a misunderstanding – as stated in the caption, Fig 1B shows Gaussian curves of error weighted average ZHe ages. These error weighted averages/Gaussian curves are based on the measured data as reported in Figure 2, Table 1 and Supplementary Data 3, and they are not related to Bayesian analysis as assumed by the reviewer. Thus, no correction is needed in our opinion.

  Fig. 2 shows histograms of posterior distributions for credible mean values for ZHe and Ar ages that are permissible by the measured data. It is important to realize that these are not the measured data (or the data set) as assumed by the reviewer. These are histograms of credible parameters (i.e., mean values and difference of the means), which are "allowable" by the measured data. This is achieved by Markov chain Monte Carlo method (MCMC) method which computes these "permissible/allowable/credible" parameters (in this case means) by trying 1,000,000 iterations and testing whether these acceptable by the measured data. In any case, obviously this was not clear to the reviewer, so we modified the figure caption accordingly and now we provide a more detailed explanation of the figure content.

  With regard to the Bayesian analysis applied in this study – we provide a complete explanation of the approach in the Methods section, which (admittedly) comes after the main story of the paper and may feel a bit out-of-sequence. However, this seems to be the format of the journal. For statistics enthusiasts we provide an in-depth explanation of the approach plus an R-code in the Supplementary Info 3 and 4.

This would provide a transparent look into support for the claim of a significant age difference between the ZHe dates from 3 domes, and the Ar and ZHe dates from the YTT samples.

- agreed, the figure is added to the main text; the distributions are now better explained in figure captions.

Editor's comments:

- move Supplementary Figure 1 into the main body (as suggested by reviewer 3) along with an explanation of how Bayesian statistical methods yield the very simple looking normal distributions shown in Fig. 1B and Fig. 2.

  - Supplementary Figure 1 was moved to the main body as the new Figure 2;
  - the very simple looking normal distributions in Fig. 1B are not related to Bayesian statistical methods, but represent Gaussian curves of weighted averages calculated form the measured data – this is stated in the Fig. 1B caption;
  - the very simple looking normal distributions in Fig. 2 (now Figure 3) are credible values for means (and mean differences) permissible by the measured data and calculated by Bayesian inference using Markov chain Monte Carlo method (MCMC). This is now better explained in the Figure 3 caption and detailed in the Methods section and in Supplementary Info 3. (Please see also our explanation above).
  -

- expand discussion of radiometric data and discuss implications for the claim of a significant age difference between the ZHe dates from 3 domes, and the Ar and ZHe dates from the YTT samples.

  - We do not fully understand this comment as we had already done the former, and all four subchapters in our Results and Discussion discuss the implications of the difference in the ages on the pre-eruptive evolution of Toba. The difference between the ZHe and Ar data is resolved (and quantified) by Bayesian analysis. As stated in our Results and shown in Figure 3, Bayesian analysis of currently available data shows that the Ar and ZHe ages from the domes are credibly (with 100% probability) different (i.e., we can be 100% sure that they are different), so we are not sure what else needs to be discussed with this regard. Nevertheless, we added a few sentences explaining the "credible difference" and probability of the difference of the means (L.131-138). We also explain and discuss the potential of Bayesian inference for geochronology in section "Bayesian inference for comparing two groups" and re-state it in our Conclusions.
  - We have tried to further clarify the implications of our findings particularly in terms of caldera-evolution and eruption dynamics. The abstract has been edited to bring our interpretations and their implications to the fore, and a few places in the text have been augmented.

- check whether the captions of Supplementary Figures 4 and 5 are swapped, and correct this if necessary.

  - corrected, captions were swapped, thank you for spotting this (both captions are very similar hence the confusion)