Authors' Response to Reviewer 2 of

A statistical analysis of zircon age distributions in volcanic, porphyry and plutonic rocks

Chetan Nathwani, Dawid Szymanowski, Lorenzo Tavazzani, Sava Markovic, Adrianna L. Virmond, and Cyril Chelle-Michou

Geochronology Discussions, https://doi.org/10.5194/gchron-2024-25

RC: Reviewers' Comment, **AR**: Authors' Response

We thank the reviewer for their supportive and constructive feedback on our manuscript and we are pleased they find it to be useful contribution to the geochronology community. They provide a series of very useful minor comments on the manuscript which we respond to individually below.

General Comments:

This manuscript addresses a critical issue in modern igneous petrology: How best to understand the measured timescales of crystallization in igneous rocks. The lens here is an important subset of this problem, zircon crystallization timescales, when measured using whole grain analyses at precisions that are much higher than the crystallization timescales.

In this work, the authors produce a large number of carefully curated U-Pb datasets, present a quantitative model for comparing their distributions (Wasserstein difference) and reducing the degrees of freedom in them (PCA), extract trends in the data that correlate broadly with the type of igneous system, and then compare these trends with predictions they make based on simple forward models. I found it to be a very clearly written paper, with good explanations of the quantitative techniques and explanations of how they related to the underlying datasets. The figures are high quality, the writing is excellent, and the choice of topic is well within the remit of Geochronology.

Specific Comments:

Like I said above, this is a very good paper. The only real issue is that the authors don't address that there are important differences between their forward models and real processes that drive crystallization and preservation of crystallization age distributions. As I'm writing this I note that this is something that the first reviewer also pointed out, so I won't spend a lot of time on it, but I think this is probably worth addressing qualitatively, for completeness. Both in terms of what the T-t history is like, and additional complications such as those highlighted by Klein and Eddy, and biases in the measured zircon record. I appreciate that it may not be possible to include those complications at this time in a quantitative model, but they are both real and important.

We have incorporated more of a discussion on the differences between our modelling and real processes, including much more detail on what could explain the difference between the results of our statistical analysis and thermal models from other studies. A more detailed response can be found in our reply to Reviewer 1. We have also made the reader aware of the caveat of our study that it does not consider some of the complications highlighted by Klein and Eddy (see below).

Specific comments/technical corrections

L43-44: It's worth noting that sampling biases may also play a role. I can't speak for every analyst, of course, but typically the largest and highest quality crystals are selected from a population. The size bias alone may be significant, as is the fact that whole grains modified by CA bias U-Pb dates by U concentration, dissolve components that are accessible to HF during the partial dissolution phase of CA, and integrate the remaining grain.

We agree that it would be useful to make a note of some additional biases. We have thought not to get into too much detail about human biases invoked by ID-TIMS as this is something that will not be unique to a specific type of sample, or the chemical abrasion biases as we do not want to convolute and confuse the paper. We do think however that it is worth noting the fact that there is a bias of bulk grain analyses to younger ages due to the volumetric differences in core to rim, as pointed out by Klein and Eddy (2023. We specify therefore that our model assumes instantaneous crystallisation of zircon

L120: This is an excellent summary of the method.

Thank you.

L134-137: I think it's fair to stick to one type of dataset but the objections to 230/238 geochronology aren't strong. Variable and asymmetric uncertainties aren't difficult to deal with, and initial [230/238] are determined in many high quality datasets (and doesn't vary infinitely). There are good reasons that plutonic datasets are rare (or absent?) and that's because the technique is exclusively applicable to young rocks, and that's a reason enough alone to exclude it.

This is fair. By far the most important reason to not include these data is their being limited to volcanic samples (simply because there are few if any plutons this young and already exposed). The asymmetric uncertainties are not a big deal, but the lack of a robust way to estimate melt-(230/238) corresponding to each zircon date remains a major source of uncertainty. We have adjusted this part to reflect this in the revised manuscript.

L159: This is very good.

Thanks!

L201-206: I'm a bit confused by this paragraph and unsure as to what I'm supposed to take away from it as a reader. It describes a technique that is not used, justifies why it is suboptimal, but then in the last sentence tells me that the results are the same? I have no specific objection here but I think the authors might want to clarify for a reader what they should take away from this. For me personally, as a geochemist who has used MDS in the past, I'd be just as happy if this paragraph didn't exist? I don't think many readers will look at this and wonder why they didn't use MDS.

This is a fair point that it does not necessarily add anything to the manuscript and the reader will not wonder why the readers did not use MDS. We therefore have accepted the reviewer's suggestion and removed this from the revised manuscript.

L220-221: Minor point but mixing and matching confidence/coverage intervals is confusing to a reader (the parameter fit uncertainties are at a coverage factor of 1 and the result is at 2).

Yes, this is a valid point thanks, the standard errors on the fit parameters have now been updated to 2SE in line with the 2σ of the output.

L221: This is a really useful result, and I could see it being something used by others in future work – it might be worth adding a bit more context for those who do? Can you specify the time range that this is applicable to, and if the fit parameters uncertainties are correlated? In the introduction you list 0.1-4567 Ma, but that's probably not relevant here (if I do the calculation right it's 40% at 0.1 Ma.)

This is calculated over the range of dates included in the compilation we used (Markovic et al.) but only up to 1000 Ma (to keep it relevant to the maximum important for our study and to avoid the fact that older dates would use Pb/Pb). We have specified this now in the revised manuscript. We agree it can be useful if one is interested in the general development of uncertainties with age, but we would not take it too far. Specifically, this shows the mid-range uncertainties in dates from multiple labs, and it uses a relatively simple fit. So real uncertainties available to any given lab may be much better or much worse.

It's also worth noting that these uncertainties differ markedly from those listed rather optimistically on line 138 (e.g., if I did the calculation correctly, it bottoms out at 0.11% in the Neoproterozoic). I'd recommend revisiting line 138 to bring it in line with the actual uncertainties parameterized in this equation – that's more of a boilerplate ID-TIMS boast than a realistic assessment of actual typical precisions from young zircon populations, where for even low-blank labs, the precision ends up being limited by the blank isotopic composition variability.

While the value in line 138 is very much "a boilerplate ID-TIMS boast", a realistic assessment of typical precisions based on that same compilation shows that it's actually fairly close to truth for all but the very youngest samples - see Fig. 11e in Markovic et al. We will adjust line 138 to reflect that reality better.

L235-240: It's worth revisiting the issues raised by Klein and Eddy at this point. Not that they need to be incorporated but just as a caveat for the reader.

Yes agreed that it would be useful to mention their work. We have decided it would best to bring this into the Methods, where we state that our model, for simplicity, assumes the instantaneous growth of each zircon and ignores the protracted growth of each crystal and the inherent bias of bulk grain analysis to younger ages. We also add a reference to Melnik and Bindeman (2018) which discusses this.

L268: Please be cautious here in assuming that natural processes won't result in gaussian distributions. The central limit theorem applies to natural processes as well – the samples we recover average a range of disparate pieces of data and proceses, and averaged samples from any distribution are gaussian.

Very fair point. This requires some careful rewording. What we mean here is that our threshold for $\Delta T/\sigma$ is probably valid as the lowest $\Delta T/\sigma$ in our filtered datasets still exhibit skew indicating they are not dominated by gaussian uncertainty (and would thus typically show normal distributions). We have rephrased to remove the statement of non-normal distributions being dominated by geological dispersion.

L305: Can you compare this to other estimates if you are aware of any? (This isn't a leading question, I'm can't find one myself)

There are some estimates made by Caricchi et al. (2016) who compared the mean, mode and median of a zircon age distribution with different numbers of zircon sampled. When sampling a normal distribution, they show that already with 10 zircons sampled, the mean, mode and median are similar to the underlying distribution. For a skewed distribution (which is what our sampling is performed on), they show that significant improvement is made when up to 30 zircons are sampled. Their sampling is performed at 100 kyr and we are unclear how/if uncertainties are propagated onto these results. Our study of the effect of number of zircons sampled is based on comparing the entire shapes of distributions rather than mean, mode and median. Given the difference in approaches and different ages under which the modelling is performed, we decided not to make a comparison in the text here as getting into the details will distract a bit the reader from the key point.

L412: This seems like an unnecessary swipe at LA-ICPMS labs. Using data that is accurate (including accurate uncertainties) is obviously critical and there's no reason to cast aspersions on a specific technique. Peak hopping TIMS Pb data on Phoenix instruments have uncertainties on isotope ratios underestimated by 1.7x due to autocorrelation (and some other factor on any other TIMS), but that fact is equally out of place here. LA-ICPMS is an easy target because of the large number of practitioners and therefore the larger n of outliers.

We do not believe we are making an unwarranted swipe at the LA-ICP-MS technique but we think it is very important to note here that using spot analyses to reconstruct an age distribution in these young rocks may be more challenging than one might think from just looking at the reported uncertainties. Quite a lot of studies report single-spot 2σ uncertainties at about 0.5% (which is likely just considering analytical uncertainty), but when taking into account repeatability of reference material analyses (which is a good approximation for systematic uncertainties asscociated with 'matrix effects' in LA-ICP-MS analyses) this is actually within the range of 2-4% which is quite a substantial difference (Sliwinski et al., 2022). Hence, we feel it appropriate here to mention as we want to avoid people overinterpreting age distributions from LA-ICP-MS data with under-estimated uncertainties. We have replaced the word "neglected" with "do not account for" in order to soften these words for the LA-ICP-MS community.

The issue of potentially underestimated uncertainties in peak-hopping TIMS Pb analyses is not insignificant as it will affect our threshold $\Delta t/\sigma$ by reducing the resolution of some datasets. It is however hard to address for now since it is not documented to a point where we could, for example, apply a correction factor to each dataset. We hope progress in data reduction algorithms for ion counter analyses will soon catch up with the needs of the community in that regard.