Dear Pieter,

Please see below by review of Mueller et al. – "An expanded workflow for detrital rutile provenance studies: An application from the Neotethys Orogen in Anatolia".

This is a nice dataset which is used as a case study to argue for a new/expanded workflow in U-Pb detrital rutile provenance studies.

There are several issues with the paper as is:

The use of the phrases "expanded workflow" (title) or "new workflow" (section 7.2 heading). This revised workflow appears to be mainly not applying U concentration thresholding in an initial trace element session. The majority of labs nowadays (as far as I know) are not doing this, so that is not new. It is shown that it is inappropriate, but if it is only undertaken by a small subset of labs then is it all that important ? it certainly doesn't warrant inclusion in the title. e.g. in L26 "We present a new workflow that accounts for low-U rutile..." - I can show you lots of published papers that date all the rutile in the rock and do not undertake U thresholding, five from my lab alone extending back to 2019.

ii) The choice of common Pb composition – it is interesting to explore the difference between the 207Pb and 208Pb methods, but they ultimately do not show much of a difference. That is new, but maybe not that significant a result. But I like the general approach to discordance filtering.

iii) the section on the choice of initial age estimate to stick into the 207Pb correction (uncorrected age [t_initial] versus the 208Pb corrected age [t_208]) is really confusing. I am really puzzled by the large difference between the two approaches for discordant data (Fig. 6). In a 2011 Chemical Geology paper I showed that the final 207Pb-corrected age differs by < 0.05% if an initial age estimate of 1 Ma is used instead of 1 Ga, demonstrating it is not dependent on the choice of initial age after five iterations. As far as I can see after five iterations in a 207Pb correction, you have converged on the answer, regardless of the starting age estimate. So I cannot explain Fig. 6 unless only one iteration of the correction has been undertaken? If that is indeed the case (only one iteration of the correction has been done), then that entire section should be removed as the process has not yet converged on a solution.

iv) the PCA approach to exploring the data is interesting, but currently not much is made of it.

I recommend major revisions, with suggestions for improvements including:

a) scaling back on the strong statements about new workflows etc. – because there are papers out there already which identify all the rutiles on a mount (SEM-EDS or Raman), analyse all grains for U-Pb and trace elements, including Cr vs Nb discrimination and Zr-in-rutile temperatures.
b) Keep the U threshold aspect in, but shorten significantly and do not make it a key aspect of the paper as I do not think it is all that common an approach nowadays

c) Clarify the choice of initial age estimate to stick into the 207Pb correction – if an iterative approach has not been used (with at least five iterations) then I am not sure why it is included. But keep the bit on discordance filtering.
d) Make more of the PCA plot, and also of your own data.

Best wishes,

David Chew

Comments working through paper:

First 8 lines of abstract. I really think there needs to be a caveat here. Very broadly speaking (and there are exceptions), rutile is not particularly common in igneous rocks in the crust and requires reasonably high pressures to crystalize – it is mainly a metamorphic mineral (where again it requires reasonably high pressures to crystallize). It is better than zircon in recording metamorphic events in provenance studies but it too has a relatively restricted paragenesis.

L46 – I am not sure why this sentence starts with "In convergent margin settings….", as I feel it is applicable to other tectonic settings as well.

L64 – and detrital apatite. In terms of recent publications (i.e. last 5 years) I feel it is more commonly used nowadays than either detrital monazite and detrital muscovite. Please also list the geochronological system applying to the mineral – U-Pb, 40Ar/39Ar, Luf-Hf etc

Section 2.1 This section needs something on the role of pressure and composition on the stability field of igneous and metamorphic rutile, and also the well documented instability instability in the sub-greenschist to lower greenschist facies (Zack et al., 2004; Yakymchuk et al., 2017).

L124 (Challenge 1). I dispute the sentence "many detrital rutile methods first analyse trace elements then only collect U-Pb data on rutile above a given U concentration threshold (4-5 ppm).". I have reviewed quite a few studies in the last few years with detrital rutile U-Pb and trace element data in them, and I have never (as far as I remember) encountered this approach. I can see why it may have been applied historically (maybe over a decade ago), where quadrupole-ICP-MS or a slow-scanning sector field MS was used to give the trace elements and U-Pb was subsequently analysed by sector field ICP-MS. But let us talk about the last few years (i.e. what is currently happening). The amount of labs doing this now I feel is very small. A modern quadrupole such as an iCAP or Agilent 7900 can easily produce all the necessary TEs and good U-Pb data simultaneously in the same spot ablation. It may appear that I am making a big deal of this - but then L125-130 then make a big deal of this. I strongly agree it would introduce a bias and this is shown later on. But I feel that such an approach is hardly ever used nowadays and so the authors are arguing against a false premise as a rationale for this paper. I feel challenge #1 needs rewriting and the screening part removed, or convincing demonstration it is still a common approach (e.g. look at all detrital rutile studies published in the last five years and find the % that did U thresholding).

I feel this is entirely restricted to sector field labs (a subset of all data produced) and only a subset of those studies would in turn screen by U thresholding.

L184 I am confused here. “We explore using an initial date estimate from the uncorrected date (ti) and from the 208Pb-corrected date (t208).” How many iterations are you using after this initial age estimate? It doesn’t really matter what the age estimate is if it eventually converges on a solution? That is what is important. Unlike for the 208Pb correction you do not specify the amount of iterations after this initial age estimate?

L188 “Note that because the correction forces intersection with the concordia, the two dates are identical”. I wouldn’t mention this at all – you have only one age when doing a 207Pb correction - you report a 207Pb-corrected date.

L190 I would like to see more about the choice of the Pb initial and whether it is appropriate to use the Stacey and Kramers (1975) model. It is well known that the 207Pb/206Pb initial ratio of metamorphic titanite is often significantly lower (i.e. more radiogenic) than the Stacey and Kramers (1975) crustal evolution model, reflecting incorporation of radiogenic Pb from rutile, a common titanite precursor (see Essex and Gromet, 2000). But rutile replacing titanite is also seen in bedrock samples - have a look at Gumsley et al. (2023, Lithos). In their Figs 11a and 11b you have metamorphic rutile with a 207Pb/206Pb initial with 0.10 -0.12, which can be convincingly linked to breakdown of late Variscan titianite

L197 presumably this principle about minimum ages also applies to 208Pb corrected data?

L200 How many iterations are used following this initial age estimate. Five was quoted for the 208Pb correction, but the number of iterations is not quoted for the 207Pb correction, and it urgently needs to be. I found in Chew et al. (2011) that it was generally insensitive to the choice of the initial age estimate input into Stacey & Kramers after a few iterations.

Section 2.3 Why are 204Pb corrections not discussed?

L216. Please also provide pressure estimates. The pressure dependence on the stability of rutile is not getting much attention in this manuscript.

L281. Significant error here – S&K at 1000 Ma is about 0.909?

Section 5.1 I found this section really hard to assess when it came to the 207Pb correction using a starting estimate of t_initial or t_208, as the amount of iterations in the 207Pb correction calculation (as far as I could see) was not explicitly specified earlier. I would be somewhat surprised to see any significant variation after a few iterations (say five). It doesn’t matter if there is a difference after one iteration – what matters is the variation after the iterative process has been completed. For example, there is a surprising large age difference in Fig 6 for the low concordance grains between an initial age estimate using t_Initial vs an initial age estimate of t_208. If this is after five iterations, then that is a noteworthy result. If it is after one iteration, then it is in my opinion of no significance as you have yet to converge on the solution. Hence I am not sure if the starting age estimate issue is all that important and could be removed (e.g. if Fig.6 is based on one iteration), but

it is hard to assess without more information. I found Fig. 6 pretty confusing to be honest and I think the figure cpation needs more information as I am not entirely sure what was being plotted.

L299-L300 “However, the similarity in the 207Pb with t208 cumulative date distribution for the 100–40% and 40–0% groups is notable,”. I really didn’t understand this clause – sorry.

L306-308. Exactly how common is this approach nowadays? To the best of my knowledge I have never reviewed a detrital rutile U-Pb paper that does this. I think nowadays it is a fairly (or even very) uncommon approach. For this reason alone, I am not sure section 5.2 it is worth including in the manuscript, certainly not in so much detail.

Figure 8 – not clear which lines defining fields belong to which paper (Triebold vs Meinhold) on the Cr/Nb plot without reading the text – label them.

Section 6.3 – PCA
Not quite sure what the main point this paragraph is trying to tell us – it could be expanded on. It does show the Cr vs Nb plot is useful in that Cr (+ V which has similar behaviour) pulls in an opposite direction to Nb (+ Ta which has similar geochemical behaviour). So the Cr vs Nb plot does a good job of separating the fields. If you were to crudely put on mafic vs pelitic fields on the PCA plot (boundary between yellow vs green), then the Hf + Zr vectors would be roughly parallel, showing the mafic vs pelitic distinction is somewhat independent of temperature. But the plot is introduced without significant additional interpretation.

Sections 7.2
Some points below link back to substantive points made at the start of the review:
“the various Pb correction methods produce similar age spectra and do not change the final provenance interpretations” – so maybe that section should be scaled a bit as ultimately it does not appear to be that important.

“the 190 Ma population is poorly represented in the detrital zircon record” - but the counterpoint needs to be made that the 90 Ma population is very important in the detrital zircon record and not in the rutile record.

What exactly is the new workflow - not doing a U-threshold and analysing all grains including those identified by SEM-EDS? There are lots of studies already doing that – I cannot see the justification for “New workflow” in the abstract text or in the heading for section 7.2. For example, Caracciolo et al. (2022) present a large U-Pb detrital rutile (n =712) dataset (along with zircon and apatite), where all rutile grains in the heavy mineral fraction determined by Raman were analysed for U-Pb and trace elements (including Cr/Nb discrimination and Zr-in-rutile temperatures). I do not think the phrase “new workflow” is justified.